МАТРИЧНЫЙ АНАЛИЗ

Цехан О.Б.

2010

ББК 22.143я73

Аннотация.

Учебное пособие содержит изложение раздела «Матричный анализ» курса «Геометрия и алгебра» в соответствии с образовательными стандартами по специальностям «Прикладная математика» и «Экономическая кибернетика», множество примеров, приложений, вопросов и заданий для самоконтроля. Учебное пособие снабжено предметным указателем, что позволяет использовать его как справочник. Будет полезно студентам и аспирантам других специальностей, пользующихся матричными методами, а также преподавателям и практикам, применяющим методы и алгоритмы матричного анализа для решения прикладных задач.

Оглавление

Предисловие

Векторы и матрицы полезны в представлении многомерных данных при моделировании и изучении абстрактных и реальных систем (в математике, технике, экономике и т.п.), описание которых требует большого количества информации. Эту информацию удобно представлять при помощи матриц. Тогда анализ систем сводится к анализу свойств матриц. Матрицы являются основным математическим аппаратом линейной алгебры и применяются при исследовании линейных отображений векторных пространств, линейных и квадратичных форм, систем линейных уравнений. Числовые алгоритмы для множества задач используют матричную и векторную арифметику. В алгоритмах оптимизации для нахождения минимума функции, например, находят применение вектор первых производных и матрица вторых производных; при реализации методов решения дифференциальных уравнений возникают матрицы специального вида.

Настоящее пособие написано на основе курса лекций по «Матричному анализу», который автор в течение ряда лет читал студентам специальностей «Прикладная математика» и «Экономическая кибернетика» факультета математики и информатики Гродненского государственного университета имени Янки Купалы. Материал составлен в соответствии с образовательным стандартом по курсу «Геометрия и алгебра» (раздел «Матричный анализ») для специальностей «Прикладная математика» и «Экономическая кибернетика» и состоит из четырех частей.

Первая часть посвящена некоторым вопросам матричной алгебры. В первой главе «Функции от матриц» вводится определение скалярной функции от матричного аргумента и способ ее построения с помощью интерполяционного полинома Лагранжа—Сильвестра. В главе «Обращение прямоугольных матриц» рассматривается псевдообратная матрица Мура—Пенроуза, способ ее построения на основе скелетного разложения, свойства и приложения псевдообращения к решению систем линейных алгебраических уравнений. В третьей главе «Матричные уравнения» излагаются способы решения некоторых типов матричных уравнений. Рассматривается метод, основанный на сведении матричных параметров уравнений к Жордановой нормальной форме, а также метод канонизации.

Во *второй части* изучаются линейные отображения и их матрицы. Рассматриваются сопряженные пространства, сопряженные отображения, унитарные, нормальные, эрмитовы и симметрические матрицы.

Третья часть посвящена некоторым прикладным аспектам линейной алгебры. Она содержит главы «Нормы векторов и матриц» и «Локализация и возмущение собственных значений». Рассматриваются нормы векторов в конечномерных и бесконечномерных пространствах, их свойства, особенности, связь матричных и векторных норм, их приложения. Приведены наиболее употребительные нормы, устанавливается связь между различными нормами. В главе 8 изложены вопросы построения областей комплексной плоскости, содержащих все собственные значения матриц. Дается изложение теории кругов Гершгорина. Рассмотрены также оценки собственных значений матриц, элементы которых подверглись возмущениям.

В *четвертой части* рассматриваются неотрицательные, в частности, примитивные, стохастические, вполне неотрицательные и осцилляционные матрицы, их свойства, характеристики и приложения.

Каждая глава разделена на параграфы и пункты. Формулы, определения, теоремы, упражнения, замечания и примеры нумеруются двумя цифрами: первая цифра — номер главы, вторая — номер формулы или соответствующего утверждения.

Предполагается, что студенты знакомы с курсом линейной алгебры. Каждая глава предваряется перечнем знаний, которыми обучаемые должны владеть на момент начала изучения темы (**Следует повторить**), и завершается перечнем минимальных знаний, умений и навыков, которые необходимо усвоить в результате изучения темы (**Необходимо усвоить**).

В каждой главе приведен теоретический материал, содержащий основные понятия и утверждения по соответствующей теме. Даны строгие доказательства всех основных утверждений. Предпочтение отдано конструктивным доказательствам. Большинство понятий и утверждений иллюстрируются примерами, поясняющими их содержание и применение к решению задач. Несложные доказательства некоторых простейших фактов предложены в виде упражнений и могут использоваться для самостоятельной работы. В конце параграфов предложены задачи и упражнения для самостоятельного изучения, к большинству из них приведены ответы и указания по выполнению. Наиболее сложные задания, умение выполнять которые не является необходимым условием усвоения курса, помечены знаком (*). Контрольные вопросы, завершающие каждый параграф, рассчитаны на обобщение и систематизацию изученного материала.

Темы, которые могут быть опущены, если не позволяет время, выделенное на изучение курса, помечены знаком (*).

При написании книги использовались некоторые материалы из источников, приведенных в списке литературы. Соответствующие ссылки даны в начале каждой главы и по тексту. Источники, содержащие дополнительный материал по теме, помечены знаком (*).

Настоящая книга является учебным пособием по разделу «Матричный анализ» курса «Геометрия и алгебра» и предназначена для студентов вузов, обучающихся по специальностям «Прикладная математика» и «Экономическая кибернетика». Будет полезна также студентам и аспирантам других специальностей, пользующихся матричными методами, а также преподавателям и практикам, применяющим методы и алгоритмы матричного анализа для решения прикладных задач.

Учебное пособие снабжено предметным указателем, что облегчит его использование в качестве справочного. Предложенные задачи, упражнения и вопросы могут использоваться как для проведения практических занятий по разделу «Матричный анализ», так и для самостоятельной работы, а также для проведения контрольных работ.

Выражаю искреннюю благодарность рецензентам — доктору физико-математических наук, профессору П.П.Забрейко, кафедре высшей математики БГУ за ценные указания и замечания, которые содействовали улучшению рукописи пособия; доценту А.С.Ляликову за помощь в подготовке макета рукописи в редакторе LaTex. Сердечно благодарю своего мужа — В.Н.Цехана за постоянную моральную поддержку, а также подготовку рисунков данного пособия.

Указатель обозначений

$\mathbb{R}$	поле вещественных чисел
$\mathbb{C}$	поле комплексных чисел
$\mathbb{C}^n$	n-мерное комплексное векторное пространство
$\mathbb{R}^n$	n-мерное вещественное векторное пространство
x, y, z и т.д.	вектор-столбцы; $x = (x_j) \in \mathbb{C}^n(\mathbb{R}^n)$
$L(x_1, \ldots, x_n)$	линейная оболочка, натянутая на векторы $(x_1, \ldots, x_n)$
$\mathbb{C}^{m\times n}$	множество $m\times n$-матриц с элементами из $\mathbb{C}$
$\mathbb{R}^{m\times n}$	множество $m\times n$-матриц с элементами из $\mathbb{R}$
A, B и т.д.	матрицы, $A = (a_{ij}) \in \mathbb{C}^{m\times n}$ $(\mathbb{R}^{m\times n})$
$\mathrm{diag}\{a_1, .., a_n\}$	диагональная $n \times n$-матрица с элементами $a_1, .., a_n$ на диагонали
E_n	единичная $n \times n$-матрица
$C[a, b]$	линейное пространство непрерывных вещественных или комплексных функций, заданных на вещественном отрезке $[a, b]$
$\langle x, y\rangle$	скалярное произведение векторов x и y
$\langle f \mid u\rangle$	скалярное произведение вектора u и ковектора f
$\overline{a}$	комплексно-сопряженное число к a
$\overline{A}$	матрица из элементов, комплексно-сопряженных к элементам матрицы $A \in \mathbb{C}^{n\times m}$

A^T	транспонированная к матрице $A \in \mathbb{C}^{n \times m}$
A^*	сопряженная к матрице $A \in \mathbb{C}^{n \times m}$, $A^* = \bar{A}^T$
A^{-1}	обратная к невырожденной матрице $A \in \mathbb{C}^{n \times n}$
A^+	обобщенная обратная матрица Мура—Пенроуза для матрицы $A \in \mathbb{C}^{m \times n}$
$\lambda_i(A)$	собственное значение матрицы A
$\sigma_i(A)$	сингулярное число матрицы A
$\operatorname{Re} z$	действительная часть числа z, $z \in \mathbb{C}$
$\operatorname{Im} z$	мнимая часть числа z, $z \in \mathbb{C}$
$\operatorname{tr} A$	след матрицы A
$p_A(\lambda)$	характеристический полином матрицы A
$m_A(\lambda)$	минимальный полином матрицы A
$\operatorname{rank} A$	ранг матрицы A
$\det A$	определитель матрицы A
$\lvert a \rvert$	абсолютное значение числа a
$\lvert A \rvert$	матрица абсолютных значений элементов матрицы A
$M(A)$	индикаторная матрица для матрицы A
$\Gamma(A)$	ориентированный граф матрицы $A \in \mathbb{C}^{n \times n}$
$\operatorname{cond}(A)$	число обусловленности матрицы A по отношению к данной матричной норме
$\sigma(A)$	спектр матрицы A
$\rho(A)$	спектральный радиус матрицы A
$J_k(\mu)$	жорданова клетка размера $k \times k$ с собственным значением μ
dim V	размерность пространства V
$\triangleq$	равно по обозначению

$\lVert\cdot\rVert_1$	l_1-норма (октаэдрическая) на $\mathbb{C}^n$; столбцовая норма на $\mathbb{C}^{n \times n}$
$\lVert\cdot\rVert_2$	l_2-норма (евклидова норма) на $\mathbb{C}^n$; спектральная норма на $\mathbb{C}^{n \times n}$
$\lVert\cdot\rVert_\infty$	l_∞-норма (кубическая) на $\mathbb{C}^n$; строчная норма на $\mathbb{C}^{n \times n}$
$\lVert\cdot\rVert_p$	l_p-норма (норма Гёльдера) на $\mathbb{C}^n$
$\lVert\cdot\rVert_E$	евклидова норма на $\mathbb{C}^{n \times n}$
$\lVert\cdot\rVert_M$	М- норма на $\mathbb{C}^{n \times n}$
$\lVert\cdot\rVert_{l_1}$	l_1-норма на $\mathbb{C}^{n \times n}$
$\lVert x \rVert$	евклидова длина (или норма) вектора $x \in \mathbb{C}^{n \times n}$
$f[\sigma(A)]$	система значений функции $f(\lambda)$ на спектре матрицы A
f'	производная функции f
$\square$	конец доказательства

Предварительные знания, необходимые для освоения раздела «Матричный анализ»

- матрицы, операции над матрицами
- специальные типы матриц и операции с ними (блочные, диагональные)
- умножение на диагональную (блочно-диагональную) матрицу слева и справа
- вырожденная, невырожденная, обратная матрица
- векторное пространство
- евклидова длина вектора
- скалярное произведение, его свойства
- ортогональность
- функция модуль и ее свойства
- понятие сходимости последовательности
- собственные числа и собственные векторы матриц
- комплексные числа, операции над ними
- комплексная плоскость
- модуль комплексного числа, его геометрическая интерпретация
- различные формы записи комплексных чисел, связь между ними
- жорданова нормальная форма
- жорданов базис
- матрица преобразования к жорданову базису
- ортогонализация Грама—Шмидта

Часть I

МАТРИЧНАЯ АЛГЕБРА

Глава 1

Функции от матриц

Матричные функции используются во многих областях линейной алгебры, возникают в многочисленных приложениях и фундаментальных исследованиях. Для решения некоторых систем дифференциальных уравнений важно уметь находить экспоненту e^{At}. Функции от матриц являются решениями нелинейных матричных уравнений ($X = A^2,\ e^X = A$). Наиболее известная матричная функция — обратная матрица, которая является распространением скалярной функции $f(\lambda) = \lambda^{-1}$ на множество матриц.

В этой главе решается вопрос о том, как произвольной числовой функции $f(\lambda)$, заданной на некотором множестве комплексных чисел и удовлетворяющей определенным требованиям, поставить в соответствие матричную функцию $f(A)$, заданную на некотором множестве матриц, т.е. распространить функцию $f(\lambda)$ и на матричные значения аргумента.

Рекомендуемая литература:

[1]*, [3]*, [4]*, [6], [9], [12], [13]*, [16], [17], [19]

Цель изучения

Изучить понятие и свойства функций от матриц, освоить метод

построения функции от матрицы с помощью интерполяционного полинома Лагранжа—Сильвестра.

Следует повторить

- аннулирующий полином матрицы
- минимальный полином матрицы
- спектр матрицы
- правильная рациональная дробь, ее представление в виде суммы элементарных дробей
- регулярная рациональная функция
- преобразование подобия

1.1. Определение функции от матриц

Простейшими функциями от матриц являются полиномы. Если задан числовой полином относительно λ :

$$f(\lambda) = \lambda^m + a_{m-1}\lambda^{m-1} + \ldots + a_0,$$

то $f(A),\ A \in \mathbb{C}^{n\times n}$, находится непосредственной подстановкой:

$$f(A) = A^m + a_{m-1}A^{m-1} + \ldots + a_0 E.$$

Рассмотрим теперь произвольную функцию $f(\lambda)$ скалярного аргумента λ. Пусть

$$m_A(\lambda) = (\lambda - \lambda_1)^{r_1}(\lambda - \lambda_2)^{r_2}\ldots(\lambda - \lambda_s)^{r_s} \qquad (1.1)$$

— минимальный полином матрицы A степени

$$r = \deg\ m_A(\lambda) = r_1 + r_2 + \ldots + r_s.$$

Здесь $\lambda_1, \ldots, \lambda_s$ — все различные собственные значения матрицы A: $\sigma(A) = \{\lambda_1, \ldots, \lambda_s\}$.

Определение 1.1. Если для функции $f(\lambda)$ в точках λ_k $\left(k=\overline{1,s}\right)$ из (1.1) определены производные

$$f(\lambda_k),\ \dot{f}(\lambda_k),\ \ddot{f}(\lambda_k),\ldots,f^{(r_k-1)}(\lambda_k),\ (k=\overline{1,s}),\tag{1.2}$$

то говорят, что *функция* $f(\lambda)$ *определена на спектре матрицы* A, а систему чисел (1.2) называют *системой значений функции* $f(\lambda)$ *на спектре матрицы* A. Совокупность этих значений будем символически обозначать $f\left[\sigma(A)\right]$.

Если функция f не определена на спектре матрицы A, то не определено и $f(A)$.

Пример 1.1 (Система значений функции на спектре матрицы). Для матрицы

$$A=\begin{bmatrix}-1&0&1\\0&2&0\\0&0&-1\end{bmatrix}.$$

минимальный полином $m_A(\lambda)=(\lambda+1)^2(\lambda-2)$ имеет два корня: $\lambda_1=-1$ кратности $r_1=2$ и $\lambda_2=2$ кратности $r_2=1$.

Для функции $f(\lambda)=e^{\lambda}$ определены производные

$$f(\lambda_1)=e^{\lambda_1}=e^{-1}=\frac{1}{e},$$
$$\dot{f}(\lambda_1)=\lambda_1e^{\lambda_1}=-e^{-1}=-\frac{1}{e},$$
$$f(\lambda_2)=e^{\lambda_2}=e^2.$$

Значит, функция e^{λ} определена на спектре матрицы A, а система значений функции $f(\lambda)=e^{\lambda}$ на спектре матрицы A имеет вид

$$f[\sigma(A)]=\{\frac{1}{e},-\frac{1}{e},e^2\}.$$

Пример 1.2 (Функция, не определенная на спектре матрицы).

Выясним, определена ли функция $f(\lambda)=\lambda^{-1}$ на спектре матрицы

$$A=\begin{bmatrix}2&1\\-4&-2\end{bmatrix}.$$

Минимальный полином матрицы $m_A=\lambda^2$ имеет один корень $\lambda=0$ кратности 2. Значение функции $f(\lambda)=\lambda^{-1}$ не определено в точке $\lambda=0$, поэтому функция $f(\lambda)=\lambda^{-1}$ не определена на спектре матрицы A, а значит, не определена и $f(A)$.

Очевидно, что каждый полином определен на спектре любой матрицы.

Лемма 1.1. *Значения полиномов $g(\lambda)$ и $h(\lambda)$ на матрице A совпадают тогда и только тогда, когда совпадают значения этих полиномов на спектре матрицы A.*

Доказательство. Пусть два полинома $g\,(\lambda)=h\,(\lambda)$ таковы, что

$$g\,(A)=h\,(A)\,.\tag{1.3}$$

Тогда $g\,(A)-h\,(A)=0$. Следовательно, разность $d\,(\lambda)=g\,(\lambda)-h\,(\lambda)$ является аннулирующим полиномом матрицы A, а значит, делится на $m_A(\lambda)$ без остатка:

$$d\,(\lambda)=g\,(\lambda)-h\,(\lambda)=m_A\,(\lambda)\,p\,(\lambda)\,.\tag{1.4}$$

Отсюда в силу (1.1) справедливо

$$d(\lambda_k)=\dot{d}(\lambda_k)=\cdots=d^{(r_k-1)}(\lambda_k)=0,\ (k=\overline{1,s}),$$

т.е.

$$g(\lambda_k)=h(\lambda_k),\ldots,g^{(r_k-1)}(\lambda_k)=h^{(r_k-1)}(\lambda_k),\ (k=\overline{1,s}).\tag{1.5}$$

Таким образом, полиномы $g(\lambda)$ и $h(\lambda)$ имеют одинаковые значения на спектре матрицы A, если выполнено (1.3).

Обратно, если полиномы $g(\lambda)$ и $h(\lambda)$ имеют одинаковые значения на спектре матрицы A, то полином $d(\lambda) = g(\lambda) - h(\lambda)$ имеет корень λ_l кратности r_l для каждого $l = 1, \ldots, s$, поэтому согласно (1.1) делится на $m_A(\lambda)$ и, значит,

$$d(A) = 0 = g(A) - h(A)$$

и равенство (1.3) выполняется. □

Таким образом, если задана матрица A, то значения полинома $g(\lambda)$ на спектре матрицы A вполне определяют матрицу $g(A)$. Значит, все полиномы $g(\lambda)$, принимающие одни и те же значения на спектре матрицы A, имеют одно и то же матричное значение $g(A)$.

Дадим определение $f(A)$ в общем случае по тому же принципу: значения функции $f(\lambda)$ на спектре матрицы A должны полностью определять $f(A)$, т.е. все функции $f(\lambda)$, имеющие одни и те же значения на спектре матрицы A, должны иметь одно и то же матричное значение $f(A)$.

Определение 1.2. Пусть функция f определена на спектре матрицы A. Тогда

$$f(A) = g(A),$$

где $g(\lambda)$ — любой полином, принимающий на спектре матрицы A те же значения, что и $f(\lambda)$:

$$f[\sigma(A)] = g[\sigma(A)].$$

Пример 1.3 (Интерполяционный полином нильпотентной клетки Жордана). Пусть H_n — нильпотентная клетка Жордана порядка n:

$$H_n = \begin{bmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 1 \\ 0 & 0 & 0 & \ldots & 0 \end{bmatrix}.$$

Для заданной функции $f(\lambda)$ найдем полином $r(\lambda)$, такой, что $r(H) = f(H)$. Минимальный полином для H_n равен $m_H(\lambda) = \lambda^n$. Поэтому значениями f на спектре H_n будут

$$f[\sigma(H_n)] = \{f(0), f^{(1)}(0), \ldots, f^{(n-1)}(0)\}.$$

Легко видеть, что искомый полином будет иметь вид:

$$r(\lambda) = f(0) + \frac{\dot{f}(0)}{1!}\lambda + \ldots + \frac{f^{(n-1)}(0)}{(n-1)!}\lambda^{(n-1)}.$$

Упражнение 1.1. Проверить, что $r[\sigma(H_n)] = f[\sigma(H_n)]$.

Заметим, что $f(H)$ имеет смысл для любой функции $f(\lambda)$, для которой определены $f(0)$, $\dot{f}(0)$, $\ldots, f^{(n-1)}(0)$.

1.2. Интерполяционный полином Лагранжа— Сильвестра

Подход, изложенный в предыдущем пункте, основан на следующей задаче: для заданной скалярной функции $f(\lambda)$ и матрицы A найти полином $r(\lambda)$ такой, что значения $f(\lambda)$ и $r(\lambda)$ на спектре матрицы совпадают, т.е. $f[\sigma(A)] = r[\sigma(A)]$.

Разрешимость поставленной задачи, называемой *«задачей интерполяции»*, гарантирует следующая

Лемма 1.2. *Пусть заданы различные числа* $p_1, p_2, \ldots, p_s$ *и таблица из* $(k+1)s$ *произвольных чисел* a_{ij}. *Найдется полином* $r(\lambda)$, *который в каждой точке* p_i *имеет значение* a_{i0}, *а его* j*-ая производная — значение* a_{ij} $(i = \overline{1,s},\ j = \overline{1,k})$.

Доказательство этого результата можно найти в [3, стр. 149], [12, стр. 4], [2, стр. 185]. □

Если существует хотя бы один полином $g(\lambda)$, значение которого на спектре матрицы A совпадает со значениями $f(\lambda)$, то таких полиномов будет бесконечно много. Но всегда можно добиться того, что степень полинома, с помощью которого мы определяем $f(A)$, будет меньше степени минимального полинома r. Действительно, разделим $g(\lambda)$ на $m_A(\lambda)$:

$$g(\lambda) = m_A(\lambda)p(\lambda) + r(\lambda), \text{ где } \deg\, r < m.$$

Тогда

$$g(A) = m_A(A)p(A) + r(A), \text{ но } m_A(A) = 0,$$

следовательно, $g(A) = r(A)$, а значит, $f(A) = r(A)$.

Существует один и только один полином $r(\lambda)$, такой, что $f[\sigma(A)] = r[\sigma(A)]$, имеющий степень, меньшую r. Он однозначно определяется интерполяционными условиями

$$r(\lambda_k) = f(\lambda_k), \ldots, r^{(r_k-1)}(\lambda_k) = f^{(r_k-1)}(\lambda_k), (k = \overline{1,s}). \quad (1.6)$$

Определение 1.3. Полином $r(\lambda)$, определяемый интерполяционными условиями (1.6), называется *интерполяционным полиномом Лагранжа—Сильвестра* функции $f(\lambda)$ на спектре матрицы A.

Тогда можно дать следующее определение $f(A)$.

Определение 1.4. Если функция $f(\lambda)$ определена на спектре матрицы A, а $r(\lambda)$ — соответствующий интерполяционный полином Лагранжа—Сильвестра, то

$$f(A) = r(A).$$

Получим представление интерполяционного полинома Лагранжа—Сильвестра.

1. Пусть все собственные значения $\lambda_1, \lambda_2, \ldots, \lambda_n$ матрицы A различны. Тогда интерполяционные условия имеют вид $f(\lambda_i) = r(\lambda_i),\ i = \overline{1,n}$, а интерполяционный полином Лагранжа—Сильвестра имеет вид:

$$r(\lambda) = \sum_{k=1}^{n} \prod_{\substack{i=1\\ i\neq k}}^{n} \frac{\lambda - \lambda_i}{\lambda_k - \lambda_i} f(\lambda_k). \quad (1.7)$$

2. Если среди собственных значений матрицы A имеются кратные, но минимальный полином имеет только простые корни $\lambda_1, \lambda_2, \ldots, \lambda_r$, то аналогично предыдущему случаю

$$r(\lambda) = \sum_{k=1}^{r} \prod_{\substack{i=1\\ i\neq k}}^{r} \frac{\lambda - \lambda_i}{\lambda_k - \lambda_i} f(\lambda_k). \quad (1.8)$$

Упражнение 1.2. Проверить, что полиномы (1.7),(1.8) решают задачу интерполяции.

3. Пусть минимальный полином матрицы A имеет корни

$$\lambda_1, \lambda_2, \ldots, \lambda_s$$

соответственно кратности

$$r_1, \ldots, r_s, \quad r_1 + r_2 + \ldots + r_s = r.$$

Получим представление интерполяционного полинома Лагранжа—Сильвестра в этом случае.

Представим функцию $\frac{r(\lambda)}{m_A(\lambda)}$, являющуюся правильной дробью ($\deg\, r < \deg\, m_A$), в виде суммы элементарных дробей:

$$\begin{aligned}\frac{r(\lambda)}{m_A(\lambda)} &= \\ &= \sum_{k=1}^{s}\left[\frac{\alpha_{k1}}{(\lambda-\lambda_k)^{r_k}} + \frac{\alpha_{k2}}{(\lambda-\lambda_k)^{r_k-1}} + \ldots + \frac{\alpha_{kr_k}}{\lambda-\lambda_k}\right],\end{aligned} \quad (1.9)$$

где α_{kj} $\left(j=\overline{1,r_k},\ k=\overline{1,s}\right)$ — некоторые числа.

Для определения этих чисел умножим обе части последнего равенства на $(\lambda-\lambda_k)^{r_k}$ и обозначим через $m_k(\lambda)$ полином

$$m_k(\lambda)=\frac{m_A(\lambda)}{(\lambda-\lambda_k)^{r_k}}.$$

Получим

$$\frac{r(\lambda)}{m_k(\lambda)}=\alpha_{k1}+\alpha_{k2}(\lambda-\lambda_k)+\ldots+\alpha_{kr_k}(\lambda-\lambda_k)^{r_k-1}+ \\ +(\lambda-\lambda_k)^{r_k}\rho(\lambda),\ k=\overline{1,s}, \tag{1.10}$$

где $\rho(\lambda)$ — рациональная функция, регулярная при $\lambda=\lambda_k$.

Отсюда

$$\alpha_{k1}=\left[\frac{r(\lambda)}{m_k(\lambda)}\right]_{\lambda=\lambda_k}, \\ \alpha_{k2}=\left[\frac{r(\lambda)}{m_k(\lambda)}\right]'_{\lambda=\lambda_k}=r(\lambda_k)\left[\frac{1}{m_k(\lambda)}\right]'_{\lambda=\lambda_k}+ \\ +r'(\lambda_k)\left[\frac{1}{m_k(\lambda)}\right]_{\lambda=\lambda_k},\ldots,k=\overline{1,s}. \tag{1.11}$$

Формулы (1.11) показывают, что числители α_{kj} в правой части равенства (1.9) выражаются через значения многочлена $r(\lambda)$ на спектре матрицы A, а эти значения нам известны: они равны соответствующим значениям функции $f(\lambda)$ и ее производных. Поэтому

$$\alpha_{kj}=\frac{1}{(j-1)!}\left[\frac{f(\lambda)}{m_k(\lambda)}\right]^{(j-1)}_{\lambda=\lambda_k},\ldots,j=\overline{1,r_k},\ k=\overline{1,s}. \tag{1.12}$$

После того, как все α_{kj} найдены, умножим обе части равенства (1.9) на $m_A(\lambda)$. Таким образом определяем $r(\lambda)$.

Окончательно, если минимальный полином имеет кратные корни, то интерполяционный полином Лагранжа—Сильвестра имеет вид:

$$r(\lambda)=\sum_{k=1}^{s}\sum_{j=1}^{r_k}\alpha_{kj}(\lambda-\lambda_k)^{j-1}m_k(\lambda). \tag{1.13}$$

Пример 1.4 (Интерполяционный полином Лагранжа—Сильвестра). Пусть матрица A имеет минимальный полином

$$m_A(\lambda)=(\lambda-\lambda_1)^2(\lambda-\lambda_2)^3,$$

функция $f(\lambda)$ определена на спектре матрицы A.

Запишем интерполяционный полином Лагранжа—Сильвестра для вычисления $f(A)$. Поскольку минимальный полином имет кратные корни, то применим формулу (1.13). Имеем

$$s=2,\ r_1=2,\ r_2=3, \\ m_1(\lambda)=(\lambda-\lambda_2)^3, \\ m_2(\lambda)=(\lambda-\lambda_1)^2.$$

Значит,

$$r(\lambda)=(\alpha_{11}+\alpha_{12}(\lambda-\lambda_1))(\lambda-\lambda_2)^3+ \\ +\left(\alpha_{21}+\alpha_{22}(\lambda-\lambda_2)+\alpha_{23}(\lambda-\lambda_2)^2\right)(\lambda-\lambda_1)^2.$$

Найдем коэффициенты:

$$\alpha_{11}=\frac{f(\lambda_1)}{(\lambda_1-\lambda_2)^3}.$$

Чтобы найти α_{12}, необходимо найти производную функции $\left[\frac{f(\lambda)}{(\lambda-\lambda_2)^3}\right]'$:

$$\left[\frac{f(\lambda)}{(\lambda-\lambda_2)^3}\right]'\Bigg|_{\lambda=\lambda_1}=\frac{f'(\lambda)(\lambda-\lambda_2)^3-3f(\lambda)(\lambda-\lambda_2)^2}{(\lambda-\lambda_2)^6}\Bigg|_{\lambda=\lambda_1}.$$

Следовательно,

$$\alpha_{12} = \frac{f'(\lambda_1)(\lambda_1-\lambda_2) - 3f(\lambda_1)}{(\lambda_1-\lambda_2)^4}.$$

Аналогично,

$$\alpha_{21} = \frac{f(\lambda_2)}{(\lambda_1-\lambda_2)^2},$$
$$\alpha_{22} = \frac{f'(\lambda_2)(\lambda_2-\lambda_1) - 2f(\lambda_2)}{(\lambda_2-\lambda_1)^3},$$
$$\alpha_{23} = \frac{1}{2}\left[\frac{f(\lambda)}{(\lambda-\lambda_1)^2}\right]''\Bigg|_{\lambda=\lambda_2}.$$

Пример 1.5 (Функция от матрицы). Вычислим значение функции e^A, где

$$A = \begin{bmatrix} 4 & -2 \\ 6 & -3 \end{bmatrix},$$

пользуясь интерполяционным полиномом Лагранжа—Сильвестра. Характеристический полином матрицы A

$$p_A(\lambda) = \lambda(\lambda-1),$$

имеет простые корни $\lambda_1 = 0,\ \lambda_2 = 1$. Поэтому для построения интерполяционного полинома Лагранжа—Сильвестра матрицы A используем формулу (1.7):

$$r(\lambda) = -(\lambda-1) + \lambda e. \tag{1.14}$$

Подставляя в интерполяционный полином Лагранжа—Сильвестра (1.14) матрицу A, имеем:

$$e^A = r(A) = -(A-E) + Ae =$$
$$= \begin{bmatrix} -3 & 2 \\ -6 & 4 \end{bmatrix} + \begin{bmatrix} 4e & -2e \\ 6e & -3e \end{bmatrix} = \begin{bmatrix} 4e-3 & -2e+2 \\ 6e-6 & -3e+4 \end{bmatrix}.$$

Пример 1.6 (Интерполяционный полином Лагранжа—Сильвестра нильпотентной клетки Жордана). Пусть H_n — нильпотентная клетка Жордана порядка n (см. пример 1.3) и $f(\lambda)$ — некоторая функция, имеющая $(n-1)$ производных при $\lambda = 0$. Найдем интерполяционный полином Лагранжа—Сильвестра.

Минимальный полином для H_n равен λ^n и значениями f на спектре H_n будут поэтому $f(0), f^{(1)}(0), \ldots, f^{(n-1)}(0)$. Положим $f^{(j)} = f^{(j)}(0),\ j = 0, \ldots, n-1$. Тогда согласно формуле (1.13) интерполяционный полином Лагранжа—Сильвестра для функции $f(\lambda)$ на спектре матрицы A равен

$$r(\lambda) = \sum_{i=1}^{n} \frac{f^{(i-1)}}{(i-1)!}\lambda^{i-1},$$

что совпадает с результатом примера 1.3. Поэтому

$$f(H_n) = \begin{bmatrix} f^{(0)} & \frac{1}{1!}f^{(1)} & \ldots & \frac{1}{(n-1)!}f^{(n-1)} \\ 0 & f^{(0)} & \ldots & \frac{1}{(n-2)!}f^{(n-2)} \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \frac{1}{1!}f^{(1)} \\ 0 & 0 & \ldots & f^{(0)} \end{bmatrix}. \tag{1.15}$$

В частном случае, при $n = 4$:

$$f(H_4) = \begin{bmatrix} f^{(0)} & f'(0) & \frac{1}{2}f''(0) & \frac{1}{6}f'''(0) \\ 0 & f^{(0)} & f'(0) & \frac{1}{2}f''(0) \\ 0 & 0 & f^{(0)} & f'(0) \\ 0 & 0 & 0 & f(0) \end{bmatrix}.$$

Найдем $\cos H$:

$$\begin{aligned} f(\lambda) &= \cos\lambda, \\ f'(\lambda) &= -\sin\lambda, \\ f''(\lambda) &= -\cos\lambda, \\ f'''(\lambda) &= \sin\lambda, \end{aligned}$$

$$cos\, H = \begin{bmatrix} 1 & 0 & -\frac{1}{2} & 0 \\ 0 & 1 & 0 & -\frac{1}{2} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Найдем e^H :

$$e^H = \begin{bmatrix} 1 & 1 & -\frac{1}{2} & \frac{1}{6} \\ 0 & 1 & 1 & \frac{1}{2} \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

1.3. Свойства функции матрицы

1^0. Если $\lambda_1, \lambda_2, \ldots, \lambda_n$ — собственные значения матрицы n-го порядка A, то $f(\lambda_1), f(\lambda_2), \ldots, f(\lambda_n)$ — полная система собственных значений матрицы $f(A)$.

Доказательство. По определению $f(A) = r(A)$. Обозначим $\lambda_1, \lambda_2, \ldots, \lambda_n$ — все собственные значения матрицы A (с учетом их кратностей). Тогда

$$p_A(\lambda) = \det[\lambda E - A] = (\lambda - \lambda_1)(\lambda - \lambda_2)\ldots(\lambda - \lambda_n) \qquad (1.16)$$

— характеристический полином матрицы A. Разложим интерполяционный полином $r(\mu)$ на линейные множители:

$$r(\mu) = a_0(\mu - \mu_1)(\mu - \mu_2)\ldots(\mu - \mu_l). \qquad (1.17)$$

Подставим в обе части (1.17) вместо μ матрицу A:

$$r(A) = a_0(A - \mu_1 E)(A - \mu_2 E)\ldots(A - \mu_l E). \qquad (1.18)$$

Перейдем к определителям в обеих частях равенства (1.18):

$$\begin{aligned} &\det(r(A)) = a_0^n \det(A - \mu_1 E)\ldots\det(A - \mu_l E) \overset{(1.16)}{=} \\ &= a_0^n(-1)^{nl} p_A(\mu_1)\, p_A(\mu_2)\ldots p_A(\mu_l) = \\ &= a_0^n(-1)^{nl}\prod_{i=1}^{l}\prod_{k=1}^{n}(\mu_i - \lambda_k) \overset{(1.17)}{=} \\ &= r(\lambda_1)\, r(\lambda_2)\ldots r(\lambda_n). \end{aligned}$$

Заметим, что полученное равенство

$$\det[r(A)] = r(\lambda_1)\, r(\lambda_2)\ldots r(\lambda_n)$$

справедливо для любого полинома $r(\mu)$. Заменим в последнем равенстве полином $r(\mu)$ на $\lambda - r(\mu)$, где λ — некоторый параметр. Имеем:

$$\det[\lambda E - r(A)] = (\lambda - r(\lambda_1))(\lambda - r(\lambda_2))\ldots(\lambda - r(\lambda_n)),$$

что с учетом $f(A) = r(A)$ и доказывает свойство 1^0. □

2^0. Если две матрицы A и B подобны и матрица S преобразует A в B: $B = S^{-1}AS$, то матрицы $f(A)$ и $f(B)$ подобны и та же матрица S преобразует $f(A)$ в $f(B)$: $f(B) = S^{-1}f(A)S$.

Доказательство. Действительно, подобные матрицы имеют одинаковые минимальные полиномы и, следовательно, $f[\sigma(A)] = f[\sigma(B)]$. Поэтому существует интерполяционный полином $r(\lambda)$ такой, что $f(A) = r(A)$, $f(B) = r(B)$. Из $B = S^{-1}AS$ следует $B^k = S^{-1}A^kS$ $(k = 0, 1, \ldots)$, а значит, $r(B) = S^{-1}r(A)S$ и $f(B) = S^{-1}f(A)S$. □

3^0. Если A — блочно-диагональная матрица:

$$A = \operatorname{diag}\{A_1, A_2, \ldots, A_u\},$$

то

$$f(A) = \operatorname{diag}\{f(A_1), f(A_2), \ldots, f(A_u)\}. \qquad (1.19)$$

Доказательство. Для доказательства обозначим $r(\lambda)$ интерполяционный полином Лагранжа—Сильвестра для функции $f(\lambda)$ на спектре матрицы A. Тогда, как легко видеть,

$$f(A) = r(A) = \operatorname{diag}\{r(A_1), r(A_2), \ldots, r(A_u)\}. \tag{1.20}$$

С другой стороны, минимальный полином $m_A(\lambda)$ является аннулирующим полиномом для каждой из матриц $A_1, A_2, \ldots, A_u$. Поэтому из равенства $f[\sigma(A)] = r[\sigma(A)]$ следует $f[\sigma(A_1)] = r[\sigma(A_1)], \ldots, f[\sigma(A_u)] = r[\sigma(A_u)]$. Поэтому $f(A_1) = r(A_1), \ldots, f(A_u) = r(A_u)$ и равенство (1.20) превращается в равенство (1.19). □

Упражнение 1.3. Доказать, что если $f(\lambda) = g(\lambda) + h(\lambda)$ и матрицы $g(A)$ и $h(A)$ существуют, то и матрица $f(A)$ существует, причем $f(A) = g(A) + h(A)$.

Упражнение 1.4. Доказать, что если $f(\lambda) = g(\lambda)h(\lambda)$ и матрицы $g(A)$ и $h(A)$ существуют, то и матрица $f(A)$ существует, причем $f(A) = g(A)h(A)$.

Задачи и упражнения

1. Найти интерполяционный полином Лагранжа—Сильвестра $r(\lambda)$ и значение $f(A)$ функции $f(\lambda)$ для жордановой клетки $J_k(\mu)$. Для каких функций $f(\lambda)$ значение $f(A)$ имеет смысл?

2. Показать, что функция $f(\lambda) = \dfrac{1}{\lambda}$ определена для всех невырожденных матриц A и только для них, причем $f(A) = A^{-1}$.

3. Вычислить следующие значения функций от матриц:

a) $A^{100},\ A = \begin{bmatrix} 0 & 2 \\ -3 & 5 \end{bmatrix}$, b) $A^{50},\ A = \begin{bmatrix} 1 & 1 \\ -1 & 3 \end{bmatrix}$,

c) $\sqrt{A},\ A = \begin{bmatrix} 3 & 1 \\ -1 & 5 \end{bmatrix}$, d) $\sqrt{A},\ A = \begin{bmatrix} 6 & 2 \\ 3 & 7 \end{bmatrix}$,

e) $e^A,\ A = \begin{bmatrix} 3 & -1 \\ 1 & 1 \end{bmatrix}$, f) $\sin A,\ A = \begin{bmatrix} \pi - 1 & 1 \\ -1 & \pi + 1 \end{bmatrix}$,

g) $e^A,\ A = \begin{bmatrix} 4 & 2 & -5 \\ 6 & 4 & -9 \\ 5 & 3 & -7 \end{bmatrix}$, h) $\ln A,\ A = \begin{bmatrix} 4 & -15 & 6 \\ 1 & -4 & 2 \\ 1 & -3 & 3 \end{bmatrix}$,

i) $\sqrt{A},\ A = \begin{bmatrix} 5 & -1 & 2 \\ 1 & 3 & 2 \\ 0 & 0 & 4 \end{bmatrix}$ j) $e^A,\ A = \begin{bmatrix} -1 & 3 & -3 \\ 2 & -1 & 2 \\ 2 & -3 & 4 \end{bmatrix}$,

k) $e^A,\ A = \begin{bmatrix} -4 & 2 & 10 \\ -4 & 3 & 7 \\ -3 & 1 & 7 \end{bmatrix}$, l) $e^A,\ A = \begin{bmatrix} 0 & \alpha \\ -\alpha & 0 \end{bmatrix}$,

m) $e^A,\ A = \begin{bmatrix} 0 & i\alpha \\ i\alpha & 0 \end{bmatrix}$, n) $\log A,\ A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$,

o)$e^A,\ \sqrt{A},\ A = \begin{bmatrix} -7 & -4 & -3 \\ 10 & 6 & 4 \\ 6 & 3 & 3 \end{bmatrix}$,

p)$\cos A,\ \sin A,\ A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & -1 & -2 & -3 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & -1 \end{bmatrix}$.

4. Вычислить $\sin \pi A$, e^A, $\cos \pi A$, A^n для матрицы

$$A = \begin{bmatrix} 1 & 4 & 2 \\ 0 & -3 & -2 \\ 0 & 4 & 3 \end{bmatrix}.$$

5. Найти все решения уравнения $X^2 = A$ для матрицы A из задания 4. Какие их этих решений будут многочленами от A?

6. Для жордановой клетки $J_n(\mu)$ найти явную формулу для $J_n^m(\mu)$.

7*. Пусть $J_n(\mu) = \mu E_n + H_n$ — жорданова клетка порядка n и $f(\lambda)$ — некоторая функция, имеющая $(n-1)$ производных при $\lambda = \mu$. Если $f^{(j)} = f^{(j)}(\mu),\ j = 0, \ldots, n-1$, то показать, что $f(J_n(\mu))$ совпадает с матрицей (1.15).

8. Определена ли функция $f(\lambda)$ на спектре матрицы A, если минимальный полином матрицы A имеет вид $m_A(\lambda)$:
   1) $f(\lambda) = \frac{1}{\sin(\lambda-3)}$, 2) $f(\lambda) = \frac{1}{\sin(\lambda+2)}$, 3) $f(\lambda) = \frac{1}{\cos(\lambda-2)}$, 4) $f(\lambda) = \frac{1}{\cos\lambda}$, 5) $f(\lambda) = \frac{1}{\lambda+\frac{\pi}{2}}$, 6) $f(\lambda) = \frac{1}{\lambda-\frac{\pi}{2}}$;

   a) $m_A = (\lambda-3)^2(\lambda+2)^3$, b) $m_A = (\lambda-\frac{\pi}{2})^2(\lambda-2)$.

9. Какие функции из задания 8 определены на спектре матрицы A из примера 1.5? Для функций, для которых это возможно, найти значение функции $f(\lambda)$ на спектре матрицы A.

10. Доказать, что если A, B — перестановочные матрицы, то $e^A e^B = e^{A+B}$.

11. Для каких матриц A уравнение $A = e^X$ разрешимо?

Вопросы для повторения

1. Что для заданной функции $f(\lambda)$ и заданной числовой матрицы A называется системой значений функции $f(\lambda)$ на спектре матрицы A?

2. Когда заданная функция $f(\lambda)$ определена на спектре матрицы A?

3. Что такое интерполяционный полином Лагранжа—Сильвестра? Для чего он строится?

4. Запишите формулы интерполяционного полинома Лагранжа—Сильвестра.

5. Как с помощью интерполяционного полинома Лагранжа—Сильвестра вычислить значение заданной функции $f(\lambda)$ на матрице A?

6. Перечислите известные вам свойства функции от матрицы.

7. Составьте алгоритм нахождения функции от матрицы с использованием интерполяционного полинома Лагранжа—Сильвестра.

Необходимо усвоить.

Основные понятия и теоремы

- значение функции на спектре матрицы
- интерполяционный полином Лагранжа—Сильвестра
- функция от матрицы

Основные навыки и умения

- вычислять значения функции на спектре матрицы
- строить интерполяционный полином Лагранжа—Сильвестра для заданной матрицы и скалярной функции
- вычислять функции от заданных матриц с использованием интерполяционного полинома Лагранжа—Сильвестра

Глава 2

Обращение прямоугольных матриц

Операция обращения определена только для квадратных невырожденных матриц. Однако во многих ситуациях целесообразно иметь обобщение этого понятия на случай вырожденных и даже не квадратных матриц. Одним из подобных обращений является обращение Мура—Пенроуза (МП-обращение), у которого, в частности, есть такое полезное свойство, как единственность. В этой главе вначале описывается скелетное разложение произвольной прямоугольной матрицы, на основе которого затем конструктивно вводится псевдообратная матрица Мура—Пенроуза, обсуждаются ее основные свойства и приложение к решению систем линейных алгебраических уравнений (СЛАУ).

Рекомендуемая литература:

[3]*, [4], [7], [9], [11], [16], [17], [19]

Цель изучения

Освоить понятие и свойства псевдообратной матрицы Мура—Пенроуза (МП-матрицы), приобрести навыки построения МП-матрицы для произвольной матрицы, ознакомиться с инстру-

ментарием исследования и решения систем линейных алгебраических уравнений с использованием МП-матрицы.

Следует повторить

- матрицы, операции над матрицами
- определитель матрицы
- ранг матрицы
- вырожденная, невырожденная, обратная матрица
- матрицы $\overline{A}$, A^T, A^*
- свойства операции *
- векторное пространство, базис
- линейная зависимость и независимость векторов
- евклидова длина вектора: $\|x\| = (x^*x)^{\frac{1}{2}}$
- СЛАУ и их свойства

2.1. Псевдообратная матрица Мура— Пенроуза

2.1.1. Скелетное разложение матрицы

Определение 2.1. Пусть $A \in \mathbb{C}^{m\times n}$, $\operatorname{rank} A = r > 0$. Представление

$$A = BC \tag{2.1}$$

где $B \in \mathbb{C}^{m\times r}$, $C \in \mathbb{C}^{r\times n}$, называется *скелетным разложением* матрицы A.

Лемма 2.1. *В скелетном разложении (2.1)*

$$\operatorname{rank} B = \operatorname{rank} C = \operatorname{rank} A = r. \tag{2.2}$$

Доказательство. Действительно, из линейной алгебры известно, что ранг произведения двух матриц не превосходит рангов сомножителей. Поэтому,

$$r = \text{rank } A = \text{rank } BC \le \min\{\text{rank } B,\ \text{rank } C\}\,. \tag{2.3}$$

Но $\operatorname{rank} B \le r$, $\operatorname{rank} C \le r$, поскольку r — один из размеров матриц B и C. Отсюда и из (2.3), следует (2.2). □

Лемма 2.2. *Для любой матрицы $A \in \mathbb{C}^{m\times n}$ существует скелетное разложение.*

Доказательство. Чтобы получить разложение (2.1), достаточно в качестве столбцов матрицы $B = [b_1, \ldots, b_r]$ взять любые r линейно-независимых столбцов матрицы A, либо любые r линейно-независимых столбцов, через которые линейно выражаются столбцы матрицы A. Тогда произвольный j-й столбец a_j, $j = \overline{1, n}$, матрицы A будет линейной комбинацией столбцов матрицы B с коэффициентами $c_{1j}, \ldots, c_{rj}$. Эти коэффициенты и образуют j-й столбец

$$c_j = \begin{bmatrix} c_{1j} \\ \vdots \\ c_{rj} \end{bmatrix}$$

матрицы C:

$$a_j = Bc_j,\ j = \overline{1, n} \quad \Leftrightarrow \quad A = BC.$$ □

Упражнение 2.1. Ввести аналогичным образом построение матриц B и C через строки матрицы C.

Замечание 2.1. Если матрица $A \in \mathbb{C}^{m\times n}$ имеет полный ранг по столбцам ($\operatorname{rank} A = n$) (или полный ранг по строкам: $\operatorname{rank} A = m$), то в качестве матрицы B удобно взять саму матрицу A (матрицу E_m), а в качестве матрицы C — матрицу E_n (матрицу A).

Пример 2.1 (Скелетное разложение).

Рассмотрим матрицу

$$A = \begin{bmatrix} 1 & -1 & 2 & 0 \\ -1 & 2 & -3 & 1 \\ 0 & 1 & -1 & 1 \end{bmatrix} \in \mathbb{R}^{3\times 4},$$

$\text{rank}A = r = 2$, и построим для нее скелетное разложение (2.1). Согласно (2.1) матрица B будет иметь размер 3×2, а матрица C — размер 2×4. Обозначим

$$a_i, b_j, c_i,\ i = \overline{1, 4}, j = \overline{1, 2},$$

— столбцы матриц A, B, C, соответственно. Матрицу $B = [b_1, b_2]$ можно составить из *первых двух* линейно-независимых столбцов матрицы A:

$$B = [b_1, b_2] = \begin{bmatrix} 1 & -1 \\ -1 & 2 \\ 0 & 1 \end{bmatrix}.$$

Найдем матрицу C. Для этого решим уравнение $BC = A$ относительно C :

$$\left[\begin{array}{rr|rrrr} 1 & -1 & 1 & -1 & 2 & 0 \\ -1 & 2 & -1 & 2 & -3 & 1 \\ 0 & 1 & 0 & 1 & -1 & 1 \end{array}\right] \sim \left[\begin{array}{rr|rrrr} 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & -1 & 1 \end{array}\right].$$

В итоге получаем:

$$C = [c_1, c_2, c_3, c_4] = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & -1 & 1 \end{bmatrix}.$$

Окончательно имеем скелетное разложение матрицы A:

$$A = \begin{bmatrix} 1 & -1 & 2 & 0 \\ -1 & 2 & -3 & 1 \\ 0 & 1 & -1 & 1 \end{bmatrix} =$$

$$= BC = \begin{bmatrix} 1 & -1 \\ -1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & -1 & 1 \end{bmatrix}.$$

Заметим, что поскольку матрица B составлена из первых двух столбцов матрицы A, то *первые два* столбца матрицы C — единичные:

$$c_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, c_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Упражнение 2.2. Построить скелетное разложение (2.1) для матрицы A из примера 2.1,
— взяв в качестве столбцов матрицы B два последних столбца матрицы A;
— используя упражнение 2.1.

Лемма 2.3. *Скелетное разложение матрицы A неединственно.*

Доказательство. Если вместо матриц B и C в (2.1) взять $B_1 = BS$, $C_1 = S^{-1}C$, где S — любая невырожденная $r \times r$-матрица, то $A = B_1C_1$ — представление того же вида (2.1). □

Лемма 2.4. *Если B и C — компоненты скелетного разложения (2.1), то матрицы B^*B и CC^* — невырожденные.*

Доказательство. Пусть x — произвольное решение уравнения

$$B^*Bx = 0. \tag{2.4}$$

Покажем, что оно может быть только нулевым. Умножим уравнение (2.4) слева на x^*:

$$x^*B^*Bx = (Bx)^*\, Bx = 0,$$

что равносильно

$$Bx = 0. \tag{2.5}$$

Согласно лемме 2.1 (2.5) — это однородная система с матрицей полного ранга по столбцам, поэтому из (2.5) следует

$$x = 0.$$

Из того, что (2.4) имеет только нулевое решение $x = 0$, вытекает, что $\det B^*B \neq 0$. Невырожденность CC^* доказывается аналогично. □

Упражнение 2.3. Доказать невырожденность CC^*.

2.1.2. Существование и единственность псевдообратной матрицы Мура— Пенроуза.

Рассмотрим матричное уравнение

$$AXA = A. \tag{2.6}$$

Если A квадратная невырожденная матрица, то это уравнение имеет единственно решение $X = A^{-1}$. Если же A произвольная прямоугольная $m \times n$-матрица, то искомое решение X имеет размеры $n \times m$, но не определяется однозначно. В общем случае уравнение (2.6) имеет бесконечное множество решений.

Определение 2.2. Матрица $A^+ \in \mathbb{C}^{n\times m}$ называется *псевдообратной* или *обобщенной обратной матрицей Мура—Пенроуза* для матрицы $A \in \mathbb{C}^{m\times n}$, если выполнены условия:

$$AA^+A = A \tag{2.7}$$

$$A^+ = UA^* = A^*V, \tag{2.8}$$

где $U \in \mathbb{C}^{n\times n}$, $\quad V \in \mathbb{C}^{m\times m}$ — некоторые матрицы.

Условие (2.8) означает, что строки (столбцы) матрицы A^+ являются линейными комбинациями строк (столбцов) матрицы A^*.

Лемма 2.5. *Для любой матрицы $A \in \mathbb{C}^{m\times n}$ верно*

$$\mathrm{tr}AA^* = \mathrm{tr}A^*A = \sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^2. \tag{2.9}$$

Доказательство. Во-первых, поскольку $A^* = \overline{A}^T$, то, используя правила умножения матриц, легко проверить, что диагональные элементы матриц AA^* и A^*A равны

$$[AA^*]_{ii} = \sum_{j=1}^{n} a_{ij}\cdot\overline{a}_{ij} = \sum_{j=1}^{n}|a_{ij}|^2,$$
$$[A^*A]_{jj} = \sum_{i=1}^{m}\overline{a}_{ij}\cdot a_{ij} = \sum_{i=1}^{m}|a_{ij}|^2. \tag{2.10}$$

Тогда по определению следа матриц с учетом (2.10) имеем:

$$\begin{aligned}\mathrm{tr}\,AA^* &= \sum_{i=1}^{m}[AA^*]_{ii} = \sum_{i}\sum_{j}|a_{ij}|^2 = \\ &= \sum_{j}\sum_{i}|a_{ij}|^2 = \sum_{j=1}^{m}[A^*A]_{jj} = \mathrm{tr}\,A^*A,\end{aligned} \tag{2.11}$$

откуда и следует справедливость (2.9). □

Следствие 2.1. *Для матрицы $A \in \mathbb{C}^{m\times n}$ любое из равенств $AA^* = 0$ или $AA^* = 0$ влечет за собой равенство $A = 0$.*

Упражнение 2.4. Доказать следствие.

Теорема 2.1. *Для любой матрицы A псевдообратная матрица Мура—Пенроуза существует, единственна и выражается по формуле*

$$A^+ = C^+B^+ = C^*\left(CC^*\right)^{-1}\left(B^*B\right)^{-1}B^*, \tag{2.12}$$

где B и C — компоненты скелетного разложения (2.1) матрицы A.

Доказательство. Докажем существование матрицы A^+. Если $A = 0$, положим $A^+ = 0$. Пусть $A \neq 0$. Рассмотрим разложение (2.1) и будем искать сначала B^+, C^+. Из определения псевдообратной матрицы имеем:

$$BB^+B = B,\ B^+ = \tilde{U}B^* \Rightarrow B\tilde{U}B^*B = B.$$

Умножим последнее равенство слева на B^*:

$$B^*B\tilde{U}B^*B = B^*B \overset{\det B^*B\neq 0}{\Rightarrow} \tilde{U} = \left(B^*B\right)^{-1}.$$

Теперь умножая последнее равенство справа на B^*, получим

$$B^+ = \left(B^*B\right)^{-1}B^*.$$

Аналогично получаем $C^+ = C^*\left(CC^*\right)^{-1}$.

Рассмотрим матрицу (2.12) и покажем, что она удовлетворяет (2.7),(2.8), т.е. является псевдообратной.

Обозначим

$$K = \left(CC^*\right)^{-1}\left(B^*B\right)^{-1}.$$

Тогда используя (2.1) и (2.12), получим:

$$AA^+A = BCC^*\left(CC^*\right)^{-1}\left(B^*B\right)^{-1}B^*BC = BC = A,$$
$$A^+ = C^*KB^* = C^*K\left(CC^*\right)^{-1}CC^*B^* = UA^*,$$
$$A^+ = C^*KB^* = C^*B^*B\left(B^*B\right)^{-1}KB^* = A^*V,$$

где
$$U = C^* K (CC^*)^{-1} C, \quad V = B (B^* B)^{-1} K B^*.$$

Докажем теперь, что для данной матрицы A не существует двух различных псевдообратных матриц A_1^+ и A_2^+. Действительно:

$$AA_1^+ A = AA_2^+ A = A,\ A_1^+ = U_1 A^* = A^* V_1,\ A_2^+ = U_2 A^* = A^* V_2,$$

откуда имеем

$$A\left(A_1^+ - A_2^+\right) A = 0,\ \left(A_1^+ - A_2^+\right) = (U_1 - U_2) A^* = A^* (V_1 - V_2).$$

Введем обозначения

$$D = A_1^+ - A_2^+,\ U = U_1 - U_2,\ V = V_1 - V_2. \tag{2.13}$$

Тогда верно

$$ADA = 0, \quad D = UA^* = A^* V.$$

Отсюда следует

$$(DA)^* DA = A^* D^* DA = A^* V^* ADA = 0,$$

откуда с учетом следствия 2.1 вытекает, что

$$DA = 0.$$

Тогда
$$DD^* = DAU^* = 0 \Rightarrow D = 0,$$

что согласно (2.13) равносильно

$$A_1^+ = A_2^+.$$

Таким образом, единственность псевдообратной матрицы, а значит и теорема 2.1, доказана. □

Теорема 2.1 дает нам способ вычисления псевдообратной матрицы A^+ по скелетному разложению матрицы A.

Пример 2.2 (Псевдообратная матрица). Для матрицы A из примера 2.1 построим псевдообратную A^+, используя ее скелетное разложение, построенное в примере 2.1, и (2.12).

$$B^* B = \begin{bmatrix} 2 & -3 \\ -3 & 6 \end{bmatrix}, \quad (B^* B)^{-1} = \begin{bmatrix} 2 & 1 \\ 1 & \frac{2}{3} \end{bmatrix},$$

$$CC^* = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix},\ (CC^*)^{-1} = \begin{bmatrix} \frac{1}{3} & 0 \\ 0 & \frac{1}{3} \end{bmatrix} = \frac{1}{3} E.$$

$$A^+ = \frac{1}{3} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & \frac{2}{3} \end{bmatrix} \begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & 1 \end{bmatrix} = \frac{1}{9} \begin{bmatrix} 3 & 0 & 3 \\ 1 & 1 & 2 \\ 2 & -1 & 1 \\ 4 & 1 & 5 \end{bmatrix}.$$

Мы доказали, что для каждой матрицы A существует одна и только одна псевдообратная матрица Мура—Пенроуза, и если матрица A представлена своим скелетным разложением (2.1), то A^+ имеет вид (2.6).

Установим некоторые свойства матрицы A^+.

Теорема 2.2 (Свойства псевдообратной матрицы Мура—Пенроуза). *Справедливы следующие свойства:*

1^0. $(A^+)^* = (A^*)^+$.

2^0. $(A^+)^+ = A$.

3^0. $(AA^+)^* = AA^+$, *т.е. матрица* AA^+ *— эрмитова (см. стр. 165).*

4^0. $(A^+ A)^* = A^+ A$, *т.е. матрица* $A^+ A$ *— эрмитова.*

5^0. $A^+ AA^+ = A^+$.

6^0. *Матрицы* A, A^+, AA^+ *и* $A^+ A$ *имеют один и тот же ранг.*

7^0. $A^+ = A^*(AA^*)^{-1}$, *если A имеет полный ранг по строкам.*

8^0. $A^+ = (A^*A)^{-1}A^*$, *если A имеет полный ранг по столбцам.*

9^0. $A = 0 \Leftrightarrow A^+ = 0$.

Доказательство. Для доказательства 1^0 используем представление (2.12). С учетом свойств операции * имеем:

$$\begin{aligned}(A^+)^* &\overset{(2.12)}{=} \left(C^*(CC^*)^{-1}(B^*B)^{-1}B^*\right)^* = \\ &= B(B^*B)^{-1}(CC^*)^{-1}C.\end{aligned} \tag{2.14}$$

Далее заметим, что для матрицы A^* представление $A^* = C^*B^*$ является скелетным разложением. Используя для A^* представление (2.12), получим

$$\begin{aligned}(A^*)^+ &= (C^*B^*)^+ = (B^*)^*(B^*(B^*)^*)^{-1}\times \\ &\times((C^*)^*C^*)^{-1}(C^*)^* = B(B^*B)^{-1}(CC^*)^{-1}C.\end{aligned} \tag{2.15}$$

Сравнивая (2.14) и (2.15), убеждаемся в справедливости 1^0.

Докажем свойство 2^0. Сформируем матрицы

$$\begin{aligned}L &= C^*(CC^*)^{-1} \in \mathbb{C}^{n\times r}, \\ K &= (B^*B)^{-1}B^* \in \mathbb{C}^{r\times m},\end{aligned} \tag{2.16}$$

где B, C — матрицы из скелетного разложения (2.1). Из представления (2.16), леммы 2.1 с учетом $\operatorname{rank} C = \operatorname{rank} C^* = \operatorname{rank} B = \operatorname{rank} B^* = r$ следует, что $\operatorname{rank} L = \operatorname{rank} K = r$. Поэтому в силу (2.12) представление

$$A^+ = LK \tag{2.17}$$

является скелетным разложением матрицы A^+.

Построим по (2.12) псевдообратную матрицу для A^+.

$$\begin{aligned}\left(A^+\right)^+ &\overset{(2.1),(2.17)}{=\!=\!=\!=\!=} K^+L^+ = K^*(KK^*)^{-1}(L^*L)^{-1}L^* = \\ &\overset{(2.16)}{=\!=\!=} B(B^*B)^{-1}\left[(B^*B)^{-1}B^*B(B^*B)^{-1}\right]^{-1}\times \\ &\times\left[(CC^*)^{-1}CC^*(CC^*)^{-1}\right]^{-1}(CC^*)^{-1}C = BC \overset{(2.1)}{=\!=} A.\end{aligned}$$

Докажем теперь свойство 3^0. Согласно (2.1), (2.12) имеем:

$$AA^+ = BCC^*(CC^*)^{-1}(B^*B)^{-1}B^* = B(B^*B)^{-1}B^*. \tag{2.18}$$

Далее с учетом свойств операции * получим

$$\begin{aligned}\left(AA^+\right)^* &= \left(B(B^*B)^{-1}B^*\right)^* = (B^*)^*\left((B^*B)^{-1}\right)^*B^* = \\ &= B((B^*B)^*)^{-1}B^* = B(B^*B)^{-1}B^*.\end{aligned} \tag{2.19}$$

Сравнивая (2.18) и (2.19), убеждаемся в справедливости 3^0.

Аналогично доказывается 4^0. Свойство 5^0 доказывается непосредственно с применением (2.12).

Докажем 6^0. Так как ранг произведения матриц не превосходит ранга любого из сомножителей, то из (2.7) и свойства 5^0 имеем:

$$\begin{aligned}\operatorname{rank} A &= \operatorname{rank} AA^+A \le \operatorname{rank} A^+, \\ \operatorname{rank} A^+ &= \operatorname{rank} A^+AA^+ \le \operatorname{rank} A.\end{aligned} \tag{2.20}$$

Отсюда следует $\operatorname{rank} A = \operatorname{rank} A^+$. Аналогично доказывается равенство двух других рангов.

7^0 и 8^0 вытекают из (2.12) с учетом замечания 2.1. □

Упражнение 2.5. Доказать подробно все свойства 1^0—9^0.

Упражнение 2.6. Пусть $A - m\times n$-матрица. Показать, что если A имеет полный ранг по строкам, то $AA^+ = E$; если A имеет полный ранг по столбцам, то $A^+A = E$.

2.2. Применение МП-матрицы для решения СЛАУ

Одним из важных свойств обращения Мура—Пенроуза является то обстоятельство, что оно позволяет найти *точное* решение системы линейных алгебраических уравнений.

Теорема 2.3. *Общее решение однородной системы*

$$Ax = 0 \tag{2.21}$$

задается равенством

$$x = \left(E - A^{+}A\right)\, q, \tag{2.22}$$

где q — произвольный вектор подходящего размера.

Доказательство. Во-первых, для любого вектора q справедливо

$$A\left[\left(E - A^{+}A\right)\, q\right] = Aq - AA^{+}Aq \overset{(2.7)}{=\!=} Aq - Aq = 0. \tag{2.23}$$

Это означает, что x — решение системы (2.21). Во-вторых, для любого решения x системы $Ax = 0$ (2.21) найдется вектор q, при котором x имеет представление (2.22). Действительно, можно просто положить $q = x$, так как

$$\left(E - A^{+}A\right) x = x - A^{+}Ax \overset{(2.21)}{=\!=} x. \tag{2.24}$$

Это заканчивает доказательство. □

Следствие 2.2. *Решение системы (2.21) единственно тогда и только тогда, когда матрица A имеет полный ранг по столбцам.*

Доказательство. Действительно, в этом случае из 8^0 вытекает $A^{+}A = E$ и из (2.22) следует $x = 0$. □

Если решение не единственно, то существует бесконечно много решений, задаваемых формулой (2.22).

Пример 2.3 (Решение однородного линейного алгебраического уравнения). Найдем решения уравнения $Ax = 0$ с матрицей

A из примера 2.1. Решение будем строить в виде (2.22). Воспользуемся построенной в примере 2.2 матрицей A^{+}. Имеем:

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \left(\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} - \frac{1}{9}\begin{bmatrix} 3 & 0 & 3 \\ 1 & 1 & 2 \\ 2 & -1 & 1 \\ 4 & 1 & 5 \end{bmatrix} \times\right.$$

$$\left.\times \begin{bmatrix} 1 & -1 & 2 & 0 \\ -1 & 2 & -3 & 1 \\ 0 & 1 & -1 & 1 \end{bmatrix}\right)^{*} \begin{bmatrix} q_1 \\ q_2 \\ q_3 \\ q_4 \end{bmatrix} = \begin{bmatrix} 2q_1 - q_3 - q_4 \\ 2q_2 + q_3 - q_4 \\ -q_1 + q_2 + q_3 \\ -q_1 - q_2 + q_4 \end{bmatrix},$$

где q_1, q_2, q_3, q_4 — произвольные параметры.

Рассмотрим теперь неоднородную систему

$$Ax = b. \tag{2.25}$$

Теорема 2.4. *Пусть заданы $m \times n$-матрица A и m-вектор b. Следующие утверждения эквивалентны:*

(a) *векторное уравнение $Ax = b$ имеет решение x,*

(b) $\operatorname{rank}[A, b] = \operatorname{rank}[A]$,

(c) $AA^{+}b = b$.

Доказательство. Эквивалентность утверждений (a) и (b) — это утверждение теоремы Кронекера—Капелли, доказанное в курсе линейной алгебры (см., например, [2, стр. 94], [8, стр. 165]). Докажем эквивалентность утверждений (a) и (c). Предположим, что уравнение $Ax = b$ совместно. Это значит, что существует вектор $\tilde{x}$ такой, что $A\tilde{x} = b$. Следовательно,

$$b = A\tilde{x} = AA^{+}A\tilde{x} = AA^{+}b,$$

то есть c) справедливо.

Пусть теперь $AA^{+}b = b$. Возьмем $\tilde{x} = A^{+}b$. Тогда

$$A\tilde{x} = AA^{+}b = b. \qquad \square$$

Теорема 2.5. *Для того чтобы векторное уравнение (2.25) имело решение, необходимо и достаточно, чтобы*

$$AA^+b = b. \tag{2.26}$$

В этом случае общее решение может быть представлено в виде

$$x = A^+b + \left(E - A^+A\right)\ q, \tag{2.27}$$

где q — произвольный вектор подходящей размерности.

Доказательство. Необходимость и достаточность выполнения (2.26) для совместности уравнения (2.12) следует из теоремы 2.4. Покажем, что общее решение задается равенством (2.27). Предположим, что имеет место (2.27) и определим вектор

$$x^0 = x - A^+b. \tag{2.28}$$

Тогда справедлива цепочка равносильностей:

$$Ax = b \overset{(2.26)}{\Leftrightarrow} Ax = AA^+b \Leftrightarrow A\left(x - A^+b\right) = 0 \overset{(2.28)}{\Leftrightarrow} Ax^0 = 0.$$

В силу теоремы 2.3 справедливо

$$x^0 = \left(E - AA^+\right)\ q,$$

что в силу (2.28) эквивалентно

$$x = A^+b + \left(E - AA^+\right)\ q. \qquad \square$$

В (2.27) первое слагаемое представляет собой частное решение неоднородного уравнения (2.25), а второе — общее решение однородного уравнения (2.21) с той же матрицей A.

Пример 2.4 (Общее решение неоднородного уравнения).
Найдем общее решение уравнения

$$\begin{bmatrix} 1 & -1 & 2 & 0 \\ -1 & 2 & -3 & 1 \\ 0 & 1 & -1 & 1 \end{bmatrix} x = \begin{bmatrix} -1 \\ 4 \\ 3 \end{bmatrix} \tag{2.29}$$

с матрицей A из примера 2.1.

Прежде всего проверим совместность уравнения (2.29). Для этого воспользуемся условием (2.26) из теоремы 2.5. Используя результат примера 2.2, имеем:

$$AA^+b =$$
$$= \begin{bmatrix} 1 & -1 & 2 & 0 \\ -1 & 2 & -3 & 1 \\ 0 & 1 & -1 & 1 \end{bmatrix} \times \frac{1}{9}\begin{bmatrix} 3 & 0 & 3 \\ 1 & 1 & 2 \\ 2 & -1 & 1 \\ 4 & 1 & 5 \end{bmatrix} \times \begin{bmatrix} -1 \\ 4 \\ 3 \end{bmatrix} =$$
$$= \frac{1}{3}\begin{bmatrix} 2 & -1 & 1 \\ -1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} \times \begin{bmatrix} -1 \\ 4 \\ 3 \end{bmatrix} = \begin{bmatrix} -1 \\ 4 \\ 3 \end{bmatrix} = b.$$

Таким образом, система (2.29) совместна. Ее решение будем строить в виде (2.27). Найдем первое слагаемое в (2.27):

$$A^+b = \frac{1}{9}\begin{bmatrix} 3 & 0 & 3 \\ 1 & 1 & 2 \\ 2 & -1 & 1 \\ 4 & 1 & 5 \end{bmatrix}\begin{bmatrix} -1 \\ 4 \\ 3 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 2 \\ 3 \\ -1 \\ 5 \end{bmatrix}.$$

Второе слагаемое в (2.27) — это общее решение однородной системы, соответствующей исследуемой неоднородной. Это решение получено в примере 2.3. Тогда общее решение неоднородной системы имеет вид:

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 2 \\ 3 \\ -1 \\ 5 \end{bmatrix} + \begin{bmatrix} 2q_1 - q_3 - q_4 \\ 2q_2 + q_3 - q_4 \\ -q_1 + q_2 + q_3 \\ -q_1 - q_2 + q_4 \end{bmatrix} =$$
$$= \begin{bmatrix} \frac{2}{3} + 2q_1 - q_3 - q_4 \\ 1 + 2q_2 + q_3 - q_4 \\ -\frac{1}{3} - q_1 + q_2 + q_3 \\ \frac{5}{3} - q_1 - q_2 + q_4 \end{bmatrix},$$

где q_1, q_2, q_3, q_4 — произвольные параметры.

Система (2.25) совместна *для любого* b, если матрица A имеет полный ранг *по строкам* (поскольку в этом случае $AA^+ = E$). Если система совместна, то ее решение *единственно*, если матрица A имеет полный ранг *по столбцам.* Очевидно, что если матрица A имеет полный ранг по строкам и полный ранг по столбцам, то она является невырожденной матрицей и единственное решение в этом случае есть $A^{-1}b$.

2.3. Нормальное псевдорешение системы линейных уравнений

Рассмотрим систему линейных алгебраических уравнений (в общем случае несовместную) в векторном виде

$$Ax = b. \tag{2.30}$$

Определение 2.3. Вектор $r = b - Ax$ называется *невязкой* вектора x.

Длина невязки $\|r\|$ характеризует, насколько вектор x близок к решению системы. Если x является решением системы, то невязка равна нулю. Если система несовместна, то невязка всегда будет ненулевой. В этом случае можно поставить задачу: найти такой вектор x^0, что величина $\|r\|^2 = \|b - Ax^0\|^2$ примет наименьшее значение. Такой подход называют *методом наименьших квадратов.*

Определение 2.4. Вектор x^0, длина невязки которого минимальна, называется *псевдорешением* системы (2.30). Псевдорешение с минимальной длиной называется *нормальным псевдорешением* системы (2.30).

Другими словами, нормальное псевдорешение системы (2.30) x^0 имеет наименьшую длину среди всех векторов, приносящих минимум длине невязки.

Теорема 2.6. *Нормальное псевдорешение системы (2.30) всегда существует, единственно и определяется по формуле*

$$x^0 = A^+b.$$

Доказательство.* Рассмотрим произвольный столбец x и представим разность $b - Ax$ в виде

$$b - Ax = \left(b - Ax^0\right) + A\left(x^0 - x\right) = u + v,$$

где $u = b - Ax^0 = b - AA^+b$, $v = A\left(x^0 - x\right)$.

Тогда

$$\begin{aligned} \|b - Ax\|^2 &= (b - Ax)^*(b - Ax) = (u + v)^*(u + v) = \\ &= u^*u + u^*v + v^*u + v^*v, \end{aligned} \tag{2.31}$$

где $\|x\|$ — евклидова длина вектора $x \in \mathbb{C} : \|x\| = (x^*x)^{\frac{1}{2}}$. Но

$$\begin{aligned} v^*u &= \left(x^0 - x\right)^* A^*\left(b - AA^+b\right) = \\ &= \left(x^0 - x\right)^*\left(A^* - A^*AA^+\right)b. \end{aligned} \tag{2.32}$$

Исходя из скелетного разложения (2.1) и представления псевдообратной матрицы (2.12), найдем

$$A^*AA^+ = C^*B^*BCC^*\left(CC^*\right)^{-1}\left(B^*B\right)^{-1}B^* = C^*B^* = A^*.$$

Поэтому из равенства (2.32) следует

$$v^*u = 0,$$

но тогда и

$$u^*v = (v^*u)^* = 0.$$

Поэтому из равенства (2.31) находим

$$\begin{aligned} \|b - Ax\|^2 &= u^*u + v^*v = \|u\|^2 + \|v\|^2 = \\ &= \|b - Ax^0\|^2 + \|A\left(x^0 - x\right)\|^2, \end{aligned} \tag{2.33}$$

следовательно, для любого x

$$\|b - Ax\|^2 \geq \|b - Ax^0\|^2,$$

а значит, x^0 — псевдорешение системы (2.30).

Докажем, что псевдорешение является нормальным. Пусть x таково, что

$$\|b - Ax\|^2 = \|b - Ax^0\|^2.$$

Тогда, согласно равенству (2.33),

$$Az = 0, \quad \text{где} \quad z = x^0 - x. \tag{2.34}$$

С другой стороны,

$$\|x\|^2 = \left(x^0 - z\right)^* \left(x^0 - z\right) = \|x^0\|^2 + \|z\|^2 - \left(x^0\right)^* z - z^* x^0. \tag{2.35}$$

Далее, вспоминая, что $A^+ = A^* V$ (см. (2.8)), получим в силу (2.34)

$$\left(x^0\right)^* z = \left(A^+ b\right)^* z \overset{(2.8)}{=\!=\!=} (A^* V b)^* z = b^* V^* A z = 0.$$

Но тогда и

$$z^* x^0 = \left(\left(x^0\right)^* z\right)^* \overset{(2.34)}{=\!=\!=} 0.$$

Поэтому из равенства (2.35) получаем

$$\|x\|^2 = \|x^0\|^2 + \|z\|^2$$

и, следовательно,

$$\|x\|^2 \geq \|x^0\|^2,$$

причем знак равенства имеет место только при $z = 0$, т.е. при $x = x^0$, где $x^0 = A^+ b$. □

Пример 2.5 (Нормальное псевдорешение). Найдем нормальное псевдорешение уравнения

$$\begin{bmatrix} 1 & -1 & 2 & 0 \\ -1 & 2 & -3 & 1 \\ 0 & 1 & -1 & 1 \end{bmatrix} x = \begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix} \tag{2.36}$$

и длину его невязки.

Прежде всего, проверим совместность уравнения (2.36). Для этого воспользуемся условием (2.26) из теоремы 2.5. Используя результат примера 2.2, имеем:

$$AA^+ b = \begin{bmatrix} 1 & -1 & 2 & 0 \\ -1 & 2 & -3 & 1 \\ 0 & 1 & -1 & 1 \end{bmatrix} \times \frac{1}{9} \begin{bmatrix} 3 & 0 & 3 \\ 1 & 1 & 2 \\ 2 & -1 & 1 \\ 4 & 1 & 5 \end{bmatrix} \times \begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix} =$$

$$= \frac{1}{3} \begin{bmatrix} 2 & -1 & 1 \\ -1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} \times \begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \\ 3 \end{bmatrix} \neq b = \begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix}.$$

Таким образом, система (2.36) несовместна.

Найдем нормальное псевдорешение системы (2.36). Согласно теореме 2.6 имеем:

$$x^0 = \begin{bmatrix} x_1^0 \\ x_2^0 \\ x_3^0 \\ x_4^0 \end{bmatrix} = A^+ b = \frac{1}{9} \begin{bmatrix} 3 & 0 & 3 \\ 1 & 1 & 2 \\ 2 & -1 & 1 \\ 4 & 1 & 5 \end{bmatrix} \begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 2 \end{bmatrix}.$$

При этом длина минимальной невязки $\|r\| = \|b - Ax^0\|$ для несовместного уравнения (2.36) равна

$$\|b - Ax^0\| = \left(\begin{bmatrix} 3 & 3 & -3 \end{bmatrix} \begin{bmatrix} 3 \\ 3 \\ -3 \end{bmatrix} \right)^{\frac{1}{2}} = \sqrt{27},$$

и среди векторов x, дающих такую величину невязки, длина вектора x^0 минимальна и равна

$$\|x^0\| = \left(1^2 + 1^2 + 0^2 + 2^2\right)^{\frac{1}{2}} = \sqrt{6}.$$

Задачи и упражнения

1. Показать, что если матрица A имеет полный ранг по столбцам, а матрица B — по строкам, то $(AB)^+ = B^+A^+$. На примере матриц
$$A = \begin{bmatrix} 1 & 2 \\ 2 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$
убедиться, что в общем случае формула обращения произведения матриц $(AB)^{-1} = B^{-1}A^{-1}$ не обобщается на псевдообратные матрицы: $(AB)^+ \neq B^+A^+$.

2. Построить a^+, где a — вектор-столбец.

3. Доказать, что $(AA^+)^+ = AA^+$ и $(A^+A)^+ = A^+A$.

4. Показать, что если матрица A является блочно-диагональной, то A^+ — тоже блочно-диагональная матрица. Например,
$$A = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix} \Leftrightarrow A^+ = \begin{bmatrix} A_1^+ & 0 \\ 0 & A_2^+ \end{bmatrix}.$$

5. Построить скелетное разложение для заданных матриц:

   (a) $\begin{bmatrix} 2i & -1+i & -2 & -1+i \\ 0 & -2 & 0 & -4-2i \\ 3+2i & 0 & -1+3i & 1-3i \end{bmatrix}$;

   (b) $\begin{bmatrix} 1 & 0 & -1 & 1 \\ -1 & 2 & 3 & 1 \\ -1 & 0 & 1 & -1 \end{bmatrix}$;

   (c) $\begin{bmatrix} 1 & -2 & -4 & 3 \\ 0 & 2 & 4 & -2 \\ -1 & 3 & 6 & -4 \end{bmatrix}$, (d) $\begin{bmatrix} 1-i & -2 \\ 3i & i+1 \\ -i & 0 \end{bmatrix}$.

6. Доказать следующие свойства для матрицы $A \in \mathbb{C}^{m\times n}$:

   (a) $A^+ = A^{-1}$ для любой невырожденной матрицы A.

   (b) $(A^T)^{-1} = (A^{-1})^T$.

   Определение 2.5. Матрица A называется *идемпотентной*, если $A^2 = A$.

   (c) Матрицы AA^+ и A^+A идемпотентны: $(AA^+)^2 = AA^+$, $(A^+A)^2 = A^+A$.

   Матрица $(E - A^+A)$ идемпотентна.

   (d) $A^+ = A$, если матрица A симметрична и идемпотентна.

   (e) $A^*(A^+)^* A^+ = A^+ = A^+(A^+)^* A^*$.

   (f) $A^*AA^+ = A^* = A^+AA^*$.

   (g) $(A^*A)^+ = A^+(A^+)^*$, $(AA^*)^+ = (A^+)^* A^+$.

   (h) $A(A^*A)^+ A^*A = A = AA^*(AA^*)^+ A$.

   (i) $A^+ = (A^*A)^+ A^* = A^*(AA^*)^+$.

   (j) $AB = 0 \Leftrightarrow B^+A^+ = 0$.

   (k) $A^+B = 0 \Leftrightarrow A^*B = 0$.

   (l) $\mathrm{rank}\,[E - A^+A] = m - \mathrm{rank}A$.

   (m) $A(E+A)^+ = (E+A^+)^+$.

   (n) Показать, что свойство d) не допускает обращения.

7. Доказать, что если матрица A симметрична, то A^+ также симметрична и $AA^+ = A^+A$.

8. Показать, что если $\det A \neq 0$, то $(AB)^+ = B^+(ABB^+)^+$.

9. Доказать, что $A^*AB = A^*C \Leftrightarrow AB = AA^+C$.

10. Доказать, что если $\det BB^* \neq 0$, то $(AB)(AB)^+ = AA^+$.

11. Показать, что $ABB^+(ABB^+)^+ = AB(AB)^+$.

12. Вычислить псевдообратную матрицу Мура—Пенроуза для заданной матрицы:

$$\text{a) } A = \begin{bmatrix} 2 & 1 & -2 \end{bmatrix}, \text{ b) } B = \begin{bmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & -2 & 3 \\ 2 & 1 & 2 & -1 \\ 1 & 2 & -2 & 4 \\ 1 & 0 & 2 & -2 \end{bmatrix},$$

$$\text{c) } C = \begin{bmatrix} -4 & -1 \\ 0 & -1 \\ 1 & -1 \\ 1 & -1 \\ 3 & -1 \end{bmatrix}, \text{ d) } D = \begin{bmatrix} -4 & 2 \\ 0 & 0 \\ 2 & -1 \\ 6 & -3 \\ 12 & -6 \\ 0 & 0 \end{bmatrix}.$$

13. Для системы $Ax = b$ с матрицей

$$A = \begin{bmatrix} -12 & -5 & -16 & 2 & 1 & -14 \\ 11 & -7 & -6 & 0 & 1 & 7 \\ 1 & 8 & 42 & -6 & -2 & 3 \\ 1 & -20 & -23 & 1 & 3 & -10 \\ 1 & -23 & -49 & 5 & 4 & -10 \\ 1 & 11 & -14 & 4 & -1 & 9 \end{bmatrix}$$

найти нормальное псевдорешение при

a) $b = \begin{bmatrix} -62 & 23 & 163 & 6 & -106 & -135 \end{bmatrix}^T$,

b) $b = \begin{bmatrix} -62.01 & 22.98 & 163 & 6 & -106 & -135 \end{bmatrix}^T$.

Является ли псевдорешение точным решением?

14. Найти нормальное псевдорешение x^0, соответствующий ему вектор невязки r и его длину для системы $Ax = b$:

a) $A = \begin{bmatrix} 1 & 2 & -1 \\ 0 & 1 & -1 \\ 0 & 1 & -1 \end{bmatrix}$, $b = \begin{bmatrix} 6 & -12 & 24 \end{bmatrix}^T$,

b) $A = \begin{bmatrix} 1 & -1 & 1 \\ -1 & 1 & -1 \\ 0 & 1 & 0 \end{bmatrix}$, $b = \begin{bmatrix} -8 & 4 & 0 \end{bmatrix}^T$,

c) $A = \begin{bmatrix} 1 & -1 & 1 \\ -1 & 0 & -1 \\ 1 & 0 & -1 \end{bmatrix}$, $b = \begin{bmatrix} 0 & 4 & 12 \end{bmatrix}^T$.

15. Найти общее решение однородной СЛАУ $Dx = 0$ с использованием псевдообратной матрицы Мура—Пенроуза. Единственно ли найденное решение? Обосновать.

a) $D = \begin{bmatrix} 1 & 1 & 5 & -1 \\ 2 & 0 & 4 & 2 \\ -1 & 2 & 4 & -5 \end{bmatrix}$, b) $D = \begin{bmatrix} 0 & 2 & 6 & -4 \\ 1 & -3 & -3 & 7 \\ 0 & 3 & 9 & -6 \end{bmatrix}$,

c) $D = \begin{bmatrix} 1 & 2 & 8 & -3 \\ -1 & 1 & 1 & -3 \\ 0 & 1 & 3 & -2 \end{bmatrix}$.

16. * Построить векторы b_1, b_2 такие, что система $Dx = b_1$ с матрицей D из задачи 15 совместна, а система $Dx = b_2$ — несовместна. Для совместной системы найти ее общее решение с использованием псевдообратной матрицы Мура—Пенроуза. Единственно ли найденное решение? Обосновать. Для несовместной системы найти ее нормальное псевдорешение. Вычислить вектор невязки и его длину.

17. Обобщить понятие «нормального псевдорешения» и теорему 2.6 на матричное уравнение $AX = B$, где A, X, B — комплексные матрицы, порядки которых ограничены лишь условием, что написанное уравнение имеет смысл.

Вопросы для повторения

1. Пусть задана матрица $A \in \mathbb{C}^{m \times n}$ ранга r и $A = BC$ — ее скелетное разложение. Каковы размерности матриц B и C? Чему равны их ранги?
2. Единственно ли скелетное разложение?
3. Как можно построить скелетное разложение заданной матрицы? Опишите процедуру.
4. Всегда ли существует псевдообратная матрица Мура—Пенроуза? Если нет, то при каких условиях существует?
5. Единственна ли псевдообратная матрица Мура—Пенроуза для произвольной матрицы? Если нет, то при каких условиях единственна?
6. Приведите формулу для нахождения псевдообратной матрицы Мура—Пенроуза через скелетное разложение заданной матрицы.
7. Чему равна псевдообратная матрица Мура—Пенроуза для невырожденной матрицы?
8. Как устроено псевдообращение числа?
9. Чему равна псевдообратная матрица для нулевой матрицы?
10. Чему равна псевдообратная матрица Мура—Пенроуза для симметричной и идемпотентной матрицы?

11. Чему равны ранги матриц A^+, AA^+ и A^+A, если $\operatorname{rank} A{=}r$?
12. К каким специальным классам матриц принадлежат матрицы AA^+, A^+A? Докажите.
13. Запишите выражение для общего решения однородной системы $Ax = 0$ с использованием псевдообратной матрицы Мура—Пенроуза.
14. Сформулируйте условия, при которых неоднородная система $Ax = b$ имеет точное решение (совместна).
15. Запишите выражение для общего решения совместной неоднородной системы $Ax = b$ с использованием псевдообратной матрицы Мура—Пенроуза.
16. Когда решение совместной системы $Ax = b$ единственно?
17. Дайте определение нормального псевдорешения (наилучшего приближенного решения) системы $Ax = b$.
18. Всегда ли система $Ax = b$ имеет нормальное псевдорешение? Если нет, то при каких условиях имеет?
19. Единственно ли нормальное псевдорешение произвольной системы $Ax = b$? Если нет, то при каких условиях единственно?
20. Запишите выражение для нормального псевдорешения системы $Ax = b$ с использованием псевдообратной матрицы Мура—Пенроуза.
21. Составьте алгоритм нахождения псевдообратной матрицы Мура—Пенроуза по компонентам ее скелетного разложения.

Необходимо усвоить.

Основные понятия и теоремы

- скелетное разложение матрицы, его свойства
- псевдообратная матрица Мура—Пенроуза

- теорема о существовании и единственности псевдообратной матрицы Мура—Пенроуза
- свойства псевдообратной матрицы Мура—Пенроуза
- представление решения СЛАУ с помощью псевдообратной матрицы Мура—Пенроуза
- нормальное псевдорешение СЛАУ
- теорема о существовании и единственности нормального псевдорешения СЛАУ

Основные умения и навыки

- строить скелетное разложение заданной матрицы
- находить псевдообратную матрицу Мура—Пенроуза для заданной марицы
- определять совместность и несовместность системы линейных алгебраических уравнений
- определять, единственно ли решение системы линейных алгебраических уравнений
- находить общее решение однородных и неоднородных совместных систем линейных алгебраических уравнений
- находить нормальное псевдорешение системы линейных алгебраических уравнений и соответствующую невязку

Глава 3

Матричные уравнения

В этой главе рассматриваются способы решения некоторых типов матричных уравнений, встречающихся в разнообразных вопросах теории матриц и ее приложений. Излагаются способы решения матричных уравнений, основанные на эквивалентных преобразованиях матриц. В первой части используется преобразование матричных коэффициентов к кононической форме Жордана, во второй — эквивалентные преобразования, обобщающие метод исключения Гаусса.

Рекомендуемая литература:

[3]*, [9], [16], [17], [19], [18]

Цель изучения

Усвоить способы построения решений некоторых типов матричных уравнений.

Следует повторить

- жорданова нормальная форма
- жорданов базис
- матрица перехода к жордановой нормальной форме

- элементарные делители числовой матрицы
- инвариантные множители числовой матрицы
- канонический базис пространства
- элементарные преобразования матриц
- элементарные матрицы
- блочно-диагональные матрицы, операции с ними

3.1. Уравнение $AX = XB$

Рассмотрим матричное уравнение

$$AX = XB \tag{3.1}$$

где A и B — заданные квадратные матрицы (вообще говоря, разных порядков), X — искомая прямоугольная матрица:

$$A \in \mathbb{C}^{m\times m}, \quad B \in \mathbb{C}^{n\times n}, \quad X \in \mathbb{C}^{m\times n}.$$

Рассмотрим элементарные делители (в поле $\mathbb{C}$) матрицы A:

$$(\lambda-\lambda_1)^{p_1}, \quad (\lambda-\lambda_2)^{p_2}, \quad \dots \quad , \quad (\lambda-\lambda_u)^{p_u}, \quad p_1+\dots+p_u=m,$$

и элементарные делители (в поле $\mathbb{C}$) матрицы B:

$$(\mu-\mu_1)^{q_1}, \quad (\mu-\mu_2)^{q_2}, \quad \dots \quad , \quad (\mu-\mu_v)^{q_v}, \quad q_1+\dots+q_v=n.$$

Через $\tilde{A}$ и $\tilde{B}$ обозначим жордановы нормальные формы матриц A и B, а через U и V — соответствующие матрицы перехода:

$$A = U\tilde{A}U^{-1}, B = V\tilde{B}V^{-1}. \tag{3.2}$$

Тогда

$$\tilde{A} = \operatorname{diag}\{J_{p_1}(\lambda_1), \dots, J_{p_u}(\lambda_u)\}, \quad p_1+p_2+\dots+p_u=m,$$

$$\tilde{B} = \operatorname{diag}\{J_{q_1}(\mu_1), \dots, J_{q_v}(\mu_v)\}, \quad q_1+q_2+\dots+q_v=n.$$

Здесь

$$J_{p_i}(\lambda_i) = \lambda_i E_{p_i} + H_{p_i} = \begin{bmatrix} \lambda_i & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda_i & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \lambda_i & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda_i \end{bmatrix}$$

— жордановы клетки матрицы A,

$$E_l = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \dots & 1 \end{bmatrix} \in \mathbb{R}^{l\times l}$$

— единичная матрица,

$$H_l = \begin{bmatrix} 0 & 1 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & \dots & 0 & 0 \end{bmatrix} \in \mathbb{C}^{l\times l}, \tag{3.3}$$

— нильпотентная клетка Жордана порядка l.

Упражнение 3.1. Проверить, что $(H_l)^l = 0$.

Подставив в (3.1) вместо A и B выражения (3.2), получим:

$$U\tilde{A}U^{-1}X = XV\tilde{B}V^{-1}. \tag{3.4}$$

Умножим обе части равенства (3.4) слева на U^{-1}, а справа — на V:

$$\tilde{A}U^{-1}XV = U^{-1}XV\tilde{B}. \tag{3.5}$$

Введя новое обозначение

$$\tilde{X} \triangleq U^{-1}XV, \ \ (\tilde{X} \in \mathbb{C}^{m\times n}), \tag{3.6}$$

запишем уравнение (3.5) так:

$$\tilde{A}\tilde{X} = \tilde{X}\tilde{B}. \tag{3.7}$$

Если нам удастся решить уравнение (3.7) относительно $\tilde{X}$, то легко находится и решение уравнения (3.1) относительно X, так как из (3.6) следует

$$X = U\tilde{X}V^{-1}. \tag{3.8}$$

Изучим структуру искомой матрицы $\tilde{X}$. В соответствии с блочно-диагональным видом матриц $\tilde{A}$ и $\tilde{B}$ матрица $\tilde{X}$ разбивается на блоки:

$$\tilde{X} = [X_{\alpha\beta}],$$

где $X_{\alpha\beta} \in \mathbb{C}^{p_\alpha \times q_\beta}$, $\alpha = \overline{1,u}$, $\beta = \overline{1,v}$.

Используя правило умножения на блочно-диагональную матрицу, выполним умножение в (3.7). Тогда это уравнение распадается на uv матричных уравнений:

$$[\lambda_\alpha E_{p_\alpha} + H_{p_\alpha}]\, X_{\alpha\beta} = X_{\alpha\beta}\left[\mu_\beta E_{q_\beta} + H_{q_\beta}\right], \quad \alpha = \overline{1,u}, \ \beta = \overline{1,v},$$

что равносильно

$$(\mu_\beta - \lambda_\alpha)\, X_{\alpha\beta} = H_{p_\alpha} X_{\alpha\beta} - X_{\alpha\beta} H_{q_\beta}. \tag{3.9}$$

Для каждого из уравнений (3.9) возможен один из двух случаев.

1. $\mu_\beta \neq \lambda_\alpha$.

Умножая обе части равенства на $\mu_\beta - \lambda_\alpha$ и заменяя согласно (3.9) $(\mu_\beta - \lambda_\alpha)\, X_{\alpha\beta}$ на $H_{p_\alpha} X_{\alpha\beta} - X_{\alpha\beta} H_{q_\beta}$, получаем:

$$\begin{aligned} &(\mu_\beta - \lambda_\alpha)^2 X_{\alpha\beta} = H_{p_\alpha}(\mu_\beta - \lambda_\alpha)\, X_{\alpha\beta} - (\mu_\beta - \lambda_\alpha)\, X_{\alpha\beta} H_{q_\beta} = \\ &= H_{p_\alpha}\left(H_{p_\alpha} X_{\alpha\beta} - X_{\alpha\beta} H_{q_\beta}\right) - \left(H_{p_\alpha} X_{\alpha\beta} - X_{\alpha\beta} H_{q_\beta}\right) H_{q_\beta} = \\ &= H_{p_\alpha}^2 X_{\alpha\beta} - 2H_{p_\alpha} X_{\alpha\beta} H_{q_\beta} + X_{\alpha\beta} H_{q_\beta}^2 = \\ &= \sum_{i+j=2} (-1)^j\, k_{ij}\, H_{p_\alpha}^i X_{\alpha\beta} H_{q_\beta}^j, \quad \text{где} \quad k_{ij} \in \mathbb{C}. \end{aligned}$$

Проделывая эту операцию $r-1$ раз, получаем:

$$(\mu_\beta - \lambda_\alpha)^r X_{\alpha\beta} = \sum_{i+j=r} (-1)^j \; k_{ij} \, H^i_{p_\alpha} X_{\alpha\beta} H^j_{q_\beta}. \tag{3.10}$$

Напомним (см. упражнение 3.1), что

$$(H_{p_\alpha})^{p_\alpha} = 0, \; \left(H_{q_\beta}\right)^{q_\beta} = 0. \tag{3.11}$$

Если $r \geq p_\alpha + q_\beta$, то выполняется хотя бы одно из соотношений

$$i \geq p_\alpha, \; j \geq q_\beta,$$

поэтому в силу (3.11) либо $H^i_{p_\alpha} = 0$, либо $H^j_{q_\beta} = 0$ и (3.10) превращается в

$$(\mu_\beta - \lambda_\alpha)^r X_{\alpha\beta} = 0. \tag{3.12}$$

Поскольку в рассматриваемом случае $\mu_\beta \neq \lambda_\alpha$, то из (3.11) следует

$$X_{\alpha\beta} = 0. \tag{3.13}$$

2. $\mu_\beta = \lambda_\alpha$.

В этом случае уравнение (3.9) принимает вид:

$$H_{p_\alpha} X_{\alpha\beta} = X_{\alpha\beta} H_{q_\beta}, \tag{3.14}$$

где матрицы $H_{p_\alpha}, H_{q_\beta}$ имеют специальную структуру (3.3): элементы первой наддиагонали равны единице, а все остальные — нулю. Учитывая это, в зависимости от значений p_α и q_β имеем:

2.1. $p_\alpha = q_\beta$. Тогда

$$X_{\alpha\beta} = \begin{bmatrix} a_1 & a_2 & a_3 & \dots & a_{p_{\alpha-1}} & a_{p_\alpha} \\ 0 & a_1 & a_2 & \dots & a_{p_{\alpha-2}} & a_{p_{\alpha-1}} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 & a_1 \end{bmatrix} = T_{p_\alpha}, \tag{3.15}$$

т.е. решением уравнения является квадратная матрица, в которой все элементы ниже главной диагонали равны нулю, элементы главной диагонали — некоторому параметру a_1, элементы первой наддиагонали — параметру a_2 и т.д.

2.2. $p_\alpha < q_\beta$. Тогда

$$X_{\alpha\beta} = \left[O_{p_\alpha \times (q_\beta - p_\alpha)} | T_{p_\alpha} \right]. \tag{3.16}$$

2.3. $p_\alpha > q_\beta$. Тогда

$$X_{\alpha\beta} = \left[\frac{T_{q_\beta}}{O_{(p_\alpha - q_\beta) \times q_\beta}} \right]. \tag{3.17}$$

Про матрицы (3.15)—(3.17) говорят, что они имеют *правильную верхнюю треугольную форму.* Количество произвольных параметров в них равно $\min\{p_\alpha, q_\beta\}$.

Пример 3.1 (Матрицы правильной верхнетреугольной формы).

1. $p_\alpha = q_\beta = 4$:

$$X_{\alpha\beta} = \begin{bmatrix} a & b & c & d \\ 0 & a & b & c \\ 0 & 0 & a & b \\ 0 & 0 & 0 & a \end{bmatrix}.$$

2. $p_\alpha = 3, \; q_\beta = 5$:

$$X_{\alpha\beta} = \left[\begin{array}{cc|ccc} 0 & 0 & a & b & c \\ 0 & 0 & 0 & a & b \\ 0 & 0 & 0 & 0 & a \end{array} \right].$$

3. $p_\alpha = 5, \; q_\beta = 3$:

$$X_{\alpha\beta} = \left[\begin{array}{ccc} a & b & c \\ 0 & a & b \\ 0 & 0 & a \\ \hline 0 & 0 & 0 \\ 0 & 0 & 0 \end{array} \right].$$

Итак, уравнение (3.14) в качестве решения имеет произвольную правильную верхнюю треугольную матрицу.

Введем обозначения:

$$d_{\alpha\beta}(\lambda) = \text{НОД}\left\{(\lambda-\lambda_\alpha)^{p_\alpha}, (\lambda-\mu_\beta)^{q_\beta}\right\},$$
$$\sigma_{\alpha\beta} = \deg d_{\alpha\beta}(\lambda).$$

Заметим, что

$$\sigma_{\alpha\beta} = \begin{cases} 0, \lambda_\alpha \neq \mu_\beta, \\ \min\{p_\alpha, q_\beta\}, \ \lambda_\alpha = \mu_\beta. \end{cases}$$

Тогда число произвольных параметров в $\tilde{X}$ (а значит, и в X) равно

$$N = \sum_{\alpha=1}^{u}\sum_{\beta=1}^{v}\sigma_{\alpha\beta}.$$

Обозначим через $X_{\tilde{A}\tilde{B}}$ решение уравнения $\tilde{A}\tilde{X} = \tilde{X}\tilde{B}$ (3.7). Тогда полученные результаты можно сформулировать следующим образом.

Теорема 3.1. *Общее решение уравнения $AX = XB$ (3.1), где*

$$A \in \mathbb{C}^{m\times m}, \quad B \in \mathbb{C}^{n\times n},$$

$$A = U\tilde{A}U^{-1}, \ \ B = V\tilde{B}V^{-1},$$

$$\tilde{A} = \operatorname{diag}\left\{J_{p_1}(\lambda_1), \ldots, J_{p_u}(\lambda_u)\right\}, \ \ p_1 + p_2 + \ldots + p_u = m,$$

$$\tilde{B} = \operatorname{diag}\left\{J_{q_1}(\mu_1), \ldots, J_{q_v}(\mu_v)\right\}, \ \ q_1 + q_2 + \ldots q_v = n,$$

может быть найдено по формуле $X = UX_{\tilde{A}\tilde{B}}V^{-1}$.

Здесь $X_{\tilde{A}\tilde{B}}$ — общее решение уравнения

$$\tilde{A}\tilde{X} = \tilde{X}\tilde{B},$$

$$X_{\tilde{A}\tilde{B}} = [X_{\alpha\beta}], \ X_{\alpha\beta} \in \mathbb{C}^{p_\alpha \times q_\beta}, \ \ \alpha = \overline{1,u}, \ \beta = \overline{1,v}.$$

Если $\mu_\beta \neq \lambda_\alpha$, то $X_{\alpha\beta} = 0$, если $\mu_\beta = \lambda_\alpha$, то $X_{\alpha\beta}$ — произвольная правильная верхняя треугольная матрица.

Матрица X зависит от N произвольных параметров c_i:

$$X = \sum_{i=1}^{N} c_i X_i,$$

где

$$N = \sum_{\alpha=1}^{u}\sum_{\beta=1}^{v}\sigma_{\alpha\beta},$$

$$\sigma_{\alpha\beta} = \deg \text{\textit{НОД}}\left\{(\lambda-\lambda_\alpha)^{p_\alpha}, (\lambda-\mu_\beta)^{q_\beta}\right\}, \alpha = \overline{1,u}, \beta = \overline{1,v}.$$

Заметим, что X_i, которое получается из X, если параметру c_i присвоить значение единица, а остальным — нуль, является частным решением уравнения (3.1). Для ненулевого решения X частные решения $X_1, \ldots, X_N$ линейно независимы и образуют фундаментальную систему решений. Действительно, если это не так, тогда существует нетривиальная линейная комбинация $\sum\limits_{i=1}^{N} c_i X_i = 0$, то есть при ненулевых значениях некоторых параметров c_i матрица X, а значит, и $X_{\tilde{A}\tilde{B}}$ равны нулю, что невозможно.

Следствие 3.1.

Если матрицы A и B не имеют одинаковых собственных значений, то уравнение $AX = XB$ (3.1) имеет только нулевое решение, то есть $X = 0$.

Пример 3.2 (Решение уравнения $AX = XB$).

Найти общее решение уравнения

$$\begin{bmatrix} 2 & 6 & -15 \\ 1 & 1 & -5 \\ 1 & 2 & -6 \end{bmatrix} \cdot X = X \cdot \begin{bmatrix} -3 & 2 \\ 1 & -2 \end{bmatrix}.$$

Найдем жорданову нормальную форму для матрицы A и матрицу перехода U. Характеристическое уравнение имеет вид

$$(\lambda+1)^3=0.$$

Значит, единственное собственное значение $\lambda{=}{-1}$ кратности 3.

Подставим собственное значение в характеристическую матрицу:

$$\operatorname{rank}\ [A-\lambda E]=\operatorname{rank}\ [A+E]=\operatorname{rank}\begin{bmatrix}3 & 6 & -15\\ 1 & 2 & -5\\ 1 & 2 & -5\end{bmatrix}=1.$$

Значит, количество векторов, составляющих базис пространства собственных векторов, соответствующих собственному значению $\lambda=-1$, равно $n-1=3-1=2$.

Пространство собственных векторов, соответствующих $\lambda=-1$, имеет вид:

$$\{5\beta-2\alpha,\alpha,\beta\,|\,\alpha,\beta\in\mathbb{C}\}\,.$$

Ищем условие, при котором у собственного вектора существует присоединенный:

$$\left[\begin{array}{ccc|c}3 & 6 & -15 & 5\beta-2\alpha\\ 1 & 2 & -5 & \alpha\\ 1 & 2 & -5 & \beta\end{array}\right].$$

Следовательно, если $\alpha=\beta$, то система будет иметь решение.

Возьмем $\alpha=1$, тогда собственный вектор $e_1=[3,1,1]$, в качестве присоединенного к нему можно взять вектор $v_2=[1,0,0]$, в качестве второго собственного вектора возьмем

$$v_2=[-2,1,0]\,.$$

Получаем жорданов базис:

$$e_1=[3,1,1]\,,$$
$$v_2=[1,0,0]\,,$$
$$v_3=[-2,1,0]\,.$$

Отсюда следует, что

$$\tilde{A}=\begin{bmatrix}-1 & 1 & 0\\ 0 & -1 & 0\\ 0 & 0 & -1\end{bmatrix},\quad U=\begin{bmatrix}3 & 1 & -2\\ 1 & 0 & 1\\ 1 & 0 & 0\end{bmatrix}.$$

Система элементарных делителей матрицы A:

$$(\lambda+1)^2\,,\quad (\lambda+1)\,.$$

Находим $\tilde{B}$ и V:

$$\det\,[B-\mu E_2]=\begin{vmatrix}-3-\mu & 2\\ 1 & -2-\mu\end{vmatrix}=\mu^2+5\mu+4=0,$$

следовательно

$$\mu_1=-1,\ \mu_2=-4.$$

Отсюда

$$\tilde{B}=\begin{bmatrix}-1 & 0\\ 0 & -4\end{bmatrix}.$$

В качестве жорданова базиса можно взять собственные векторы $e_1=\begin{bmatrix}1 & 1\end{bmatrix}$, $e_2=\begin{bmatrix}-2 & 1\end{bmatrix}$.

Система элементарных делителей матрицы B:

$$(\mu+1)\,,\ (\mu+4)\,,$$

матрица перехода к жордановой нормальной форме

$$V=\begin{bmatrix}1 & -2\\ 1 & 1\end{bmatrix},\quad V^{-1}=\frac{1}{3}\begin{bmatrix}1 & 2\\ -1 & 1\end{bmatrix}.$$

Найдем $X_{\tilde{A}\tilde{B}}$:

$$A:\quad (\lambda+1)^2\,,\ (\lambda+1)\,,\qquad B:\quad (\mu+1)\,,\ (\mu+4)\,,$$

следовательно,

$$X_{\tilde{A}\tilde{B}}=\begin{array}{cc}\begin{array}{cc}\mu_1 & \mu_2\end{array} & \\ \left[\begin{array}{c|c}a & 0\\ 0 & 0\\ \hline b & 0\end{array}\right] & \begin{array}{c}\lambda_1\\ \\ \lambda_2\end{array}\end{array},\qquad a,b\in\mathbb{C}.$$

Найдем X:

$$X = \frac{1}{3}\begin{bmatrix} 3 & 1 & -2 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} a & 0 \\ 0 & 0 \\ b & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix} =$$

$$= \frac{1}{3}\begin{bmatrix} 3a-2b & 0 \\ a+b & 0 \\ a & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix} = \tag{3.18}$$

$$= \frac{1}{3}\begin{bmatrix} 3a-2b & 6a-4b \\ a+b & a2+2b \\ a & 2a \end{bmatrix}, \ a, b \in \mathbb{C}.$$

В качестве базиса пространства решений можно взять матрицы:

$$X_1 = \begin{bmatrix} 3 & 6 \\ 1 & 2 \\ 1 & 2 \end{bmatrix}, \ X_2 = \begin{bmatrix} -2 & -4 \\ 1 & 2 \\ 0 & 0 \end{bmatrix}.$$

3.2. Уравнение $AX = XA$

Рассмотрим частный случай уравнения (3.1) при $B = A$:

$$AX = XA, \ A \in \mathbb{C}^{m\times m}. \tag{3.19}$$

Задача нахождения общего решения $X \in \mathbb{C}^{m\times m}$ уравнения (3.19) равносильна нахождению всех матриц, перестановочных с матрицей A.

Сформулируем теорему 3.1 в частном случае для уравнения (3.19).

Теорема 3.2. *Общее решение уравнения $AX = XA$ (3.19), где*

$$A \in \mathbb{C}^{m\times m}, \qquad A = U\tilde{A}U^{-1},$$

$$\tilde{A} = \operatorname{diag}\{J_{p_1}(\lambda_1), \ldots, J_{p_u}(\lambda_u)\}, \ \ p_1 + \ldots + p_u = m,$$

может быть найдено по формуле $X = UX_{\tilde{A}}U^{-1}$.

Здесь $X_{\tilde{A}}$ — общее решение уравнения

$$\tilde{A}\tilde{X} = \tilde{X}\tilde{A},$$

$$X_{\tilde{A}} = [X_{\alpha\beta}], \qquad \alpha, \beta = \overline{1, u}, \ .$$

Если $\lambda_\alpha \neq \lambda_\beta$, то $X_{\alpha\beta} = 0$, если $\lambda_\alpha = \lambda_\beta$, то $X_{\alpha\beta}$ — произвольная правильная верхняя треугольная матрица. Матрица X зависит от N произвольных параметров, $N = \sum\limits_{\alpha,\beta=1}^{u} \sigma_{\alpha\beta}$, где

$$\sigma_{\alpha\beta} = \deg \text{\textit{НОД}}\{(\lambda-\lambda_\alpha)^{p_\alpha}, (\lambda-\mu_\beta)^{p_\beta}\}, \ \ \alpha, \beta = \overline{1, u}.$$

Пример 3.3 (Решение уравнения $AX = XA$). Матрица A имеет следующие элементарные делители:

$$(\lambda-\lambda_1)^4, \quad (\lambda-\lambda_2)^3, \quad (\lambda-\lambda_1)^2, \quad (\lambda-\lambda_2).$$

Тогда общее решение уравнения $\tilde{A}\tilde{X} = \tilde{X}\tilde{A}$ имеет вид:

$$X_{\tilde{A}} = \begin{array}{cc} & \begin{array}{cccc|ccc|cc|c} \multicolumn{4}{c}{\mu_1=\lambda_1} & \multicolumn{3}{c}{\mu_2=\lambda_2} & \multicolumn{2}{c}{\mu_3=\lambda_3} & \mu_4 \end{array} \\ \begin{array}{c} \\ \\ \lambda_1 \\ \\ \\ \lambda_2 \\ \\ \lambda_3 \\ \\ \lambda_4 \end{array} & \left[\begin{array}{cccc|ccc|cc|c} a & b & c & d & 0 & 0 & 0 & e & f & 0 \\ 0 & a & b & c & 0 & 0 & 0 & 0 & e & 0 \\ 0 & 0 & a & b & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & a & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & g & h & i & 0 & 0 & j \\ 0 & 0 & 0 & 0 & 0 & g & h & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & g & 0 & 0 & 0 \\ \hline 0 & 0 & k & l & 0 & 0 & 0 & m & n & 0 \\ 0 & 0 & 0 & k & 0 & 0 & 0 & 0 & m & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & p & 0 & 0 & q \end{array}\right] \end{array}$$

Рассмотрим инвариантные множители матрицы A:

$$f_1(\lambda), \ f_2(\lambda), \ \ldots, \ f_t(\lambda), \ f_{t+1} = \ldots = f_m = 1.$$

Пусть $n_j = \deg f_j(\lambda)$, тогда

$$n_1 \geq n_2 \geq \ldots > n_{t+1} = \ldots = n_m = 0.$$

Так как каждый нетривиальный инвариантный множитель является произведением нескольких попарно взаимно простых элементарных делителей, то количество параметров в решении уравнения $AX = XA$ (3.19) может быть найдено по формуле:

$$N = \sum_{j,k=1}^{t} \varepsilon_{jk},$$

где $\varepsilon_{jk} = \deg \text{НОД}\{f_1(\lambda), f_k(\lambda)\} = \min\{n_j, n_k\}$.

Отсюда мы получаем, что

$$N = n_1 + 3n_2 + 5n_3 + \ldots + (2t-1)\, n_t.$$

Таким образом, доказана следующая теорема.

Теорема 3.3. *Число линейно независимых матриц, перестановочных с матрицей $A \in \mathbb{C}^{m\times m}$, определяется формулой*

$$N = n_1 + 3n_2 + 5n_3 + \ldots + (2t-1)\, n_t,$$

где $f_1(\lambda), f_2(\lambda), \ldots, f_t(\lambda)$ — непостоянные инвариантные множители матрицы A, $n_j = \deg f_j(\lambda)$, $j = \overline{1,t}$.

Замечание 3.1. Ясно, что $m = n_1 + n_2 + \ldots + n_t$. Отсюда следует, что $N \geq m$, причем $N = m$ равносильно $t = 1$, то есть все элементарные делители матрицы A взаимно просты.

Пример 3.4 (Количество перестановочных матриц). Пусть матрица имеет следующие элементарные делители:

$$(\lambda-\lambda_1)^2, (\lambda-\lambda_1), (\lambda-\lambda_2)^4, (\lambda-\lambda_2)^3, (\lambda-\lambda_2).$$

Значит, нетривиальные инвариантные множители матрицы A имеют вид:

$$f_1(\lambda) = (\lambda-\lambda_1)^2(\lambda-\lambda_2)^4,\ n_1 = 6,$$
$$f_2(\lambda) = (\lambda-\lambda_1)(\lambda-\lambda_2)^3,\ n_2 = 4,$$
$$f_3(\lambda) = (\lambda-\lambda_2),\ n_3 = 1.$$

Тогда, согласно теореме 3.3, количество линейно независимых матриц, перестановочных с матрицей A, равно

$$N = 6 + 3\cdot 4 + 5\cdot 1 = 23.$$

3.3. Уравнение $AX - XB = C$

Пусть дано уравнение $AX - XB = C$, где

$$A \in \mathbb{C}^{m\times m},\ \ B \in \mathbb{C}^{n\times n},\ \ C \in \mathbb{C}^{m\times n},\ \ X \in \mathbb{C}^{m\times n}.$$

Это матричное уравнение эквивалентно системе $m\cdot n$ линейных уравнений относительно элементов матрицы X.

Рассмотрим соответствующее однородное уравнение

$$AX - XB = 0.$$

Из теоремы 3.1 следует, что если матрицы A и B не имеют одинаковых собственных значений, то уравнение $AX - XB = 0$ имеет только нулевое решение, а значит $AX - XB = C$ имеет единственное решение. Если же матрицы A и B имеют одинаковые собственные значения, то в зависимости от C возможны два варианта:

1. Уравнение не имеет решения.
2. $X = X_0 + X_{\tilde{A}\tilde{B}}$,

где X_0— произвольное частное решение уравнения

$$AX - XB = C,$$

а $X_{\tilde{A}\tilde{B}}$ — общее решение уравнения

$$AX - XB = 0,$$

структура которого описана в теореме 3.1.

3.4*. Метод канонизации

В этом разделе излагается метод решения матричных уравнений, основанный на приведении матричных коэффициентов к каноническим базисам.

Делители нуля

Определение 3.1. Пусть A — матрица размера $m \times n$. Говорят, что матрица R_A размера $n \times q$ — ее *правый делитель нуля*, если выполняется тождество

$$AR_A = 0_{m\times q}. \tag{3.20}$$

Если тождество (3.20) имеет место только при нулевой матрице $R_A = 0$, то говорят, что у матрицы A *правый делитель нуля отсутствует*.

Пример 3.5 (Матричные делители нуля). Для прямоугольной матрицы

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 3 & 1 & 0 & 0 \\ 0 & 4 & 2 & 0 & 0 \end{bmatrix} \tag{3.21}$$

правый делитель нуля

$$\breve{A}^R = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 \end{bmatrix}^T$$

очевиден в силу наличия нулевого столбца.

Упражнение 3.2. Показать, что правый делитель нуля отсутствует, если $\operatorname{rank} A = n$.

Определение 3.2. Для матрицы $A \in \mathbb{C}^{m\times n}$ ранга r *правым делителем нуля максимального ранга* называется матрица $\breve{A}^R \in \mathbb{C}^{n\times(n-r)}$ ранга $n-r$, которая удовлетворяет тождеству

$$A\breve{A}^R = 0_{m\times(n-r)}. \tag{3.22}$$

Определение 3.3. Для матрицы $A\in\mathbb{C}^{m\times n}$ ранга r *левым делителем нуля максимального ранга* называется матрица $\breve{A}^L \in \mathbb{C}^{(m-r)\times m}$ ранга $m-r$, которая удовлетворяет тождеству

$$\breve{A}^L A = 0_{(m-r)\times n}. \tag{3.23}$$

Левые делители нуля характеризуют все линейно зависимые комбинации строк; правые — все линейно зависимые комбинации столбцов.

Так как для любых матриц Π, Υ справедливо

$$(\Pi\breve{A}^L)A = 0, \qquad A(\breve{A}^R\Upsilon) = 0, \tag{3.24}$$

то все множество делителей нуля получается умножением любого делителя нуля на произвольную матрицу соответствующей размерности (с нужной стороны):

$$\left\{\breve{A}^L\right\}_\Pi = \Pi\,\breve{A}^L, \qquad \left\{\breve{A}^R\right\}_\Upsilon = \breve{A}^R\Upsilon, \tag{3.25}$$

где Π и Υ – произвольные матрицы подходящих размеров.

Формирование делителей нуля

Определение делителей нуля для исходной матрицы A в общем случае представляет собой непростую задачу. Однако для матриц в канонических базисах структуры делителей нуля очевидны.

Для произвольной матрицы A размера $m \times n$ ранга r существуют невырожденные матрицы $T_x \in \mathbb{C}^{n\times n}$, $T_y \in \mathbb{C}^{m\times m}$ такие, что A может быть представлена в канонических базисах в виде

$$A = T_y^{-1} \begin{bmatrix} E_r & 0_{r\times(n-r)} \\ 0_{(m-r)\times r} & 0_{(m-r)\times(n-r)} \end{bmatrix} T_x. \tag{3.26}$$

Пример 3.6 (Приведение матрицы к каноническим базисам).

Матрица (3.21) приводится к каноническим базисам преобразованием:

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 2 & 0 & -2 & 1 \end{bmatrix}}_{T_y} \underbrace{\begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 3 & 1 & 0 & 0 \\ 0 & 4 & 2 & 0 & 0 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} 0 & 1 & 0 & 0 & -1 \\ 1 & -1 & 0 & 0 & 1 \\ -2 & 2 & 1 & 0 & -2 \\ -1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}}_{T_x^{-1}} =$$

$$= \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}}_{\mathcal{A}}. \tag{3.27}$$

В случае, если матрица A прямоугольная или вырожденная, матрицы преобразования T_y и T_x^{-1} не единственны.

Матрица

$$\mathcal{A} = T_y A T_x^{-1}, \tag{3.28}$$

стоящая в (3.26) в квадратных скобках, имеет очевидные делители нуля максимального ранга вида

$$\mathcal{A}^R = \begin{bmatrix} 0_{r\times(n-r)} \\ E_{n-r} \end{bmatrix}, \tag{3.29}$$

$$\mathcal{A}^L = \begin{bmatrix} 0_{(m-r)\times r} & E_{m-r} \end{bmatrix}, \tag{3.30}$$

которые называются *каноническими*.

Непосредственными вычислениями легко убедиться, что правый делитель нуля матрицы (3.26) имеет вид

$$\breve{A}^R = T_x^{-1} \begin{bmatrix} 0_{m\times(n-m)} \\ E_{n-m} \end{bmatrix}, \tag{3.31}$$

а левый делитель нуля матрицы A (3.26) имеет вид

$$\breve{A}^L = \begin{bmatrix} 0_{(m-n)\times n} & E_{m-n} \end{bmatrix} T_y. \tag{3.32}$$

Таким образом, для нахождения делителей нуля заданной прямоугольной матрицы достаточно построить матрицы преобразований базисов $T_x \in \mathbb{C}^{n\times n}$, $T_y \in \mathbb{C}^{m\times m}$ из (3.26), а затем, по формулам (3.31) и/или (3.32) построить делители нуля.

Пример 3.7 (Делители нуля канонической матрицы). В соответствии со структурой матрицы $\mathcal{A}$ из примера 3.6 в правой части тождества (3.27) (одна нулевая строка и два нулевых столбца) матрица A из примера 3.5 имеет оба делителя нуля. Их канонический вид

$$\breve{\mathcal{A}}^L = \begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix}, \quad \breve{\mathcal{A}}^R = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Далее по (3.31) и (3.32) имеем правый и левый делители нуля:

$$\breve{A}^R = \begin{bmatrix} 0 & 1 & 0 & 0 & -1 \\ 1 & -1 & 0 & 0 & 1 \\ -2 & 2 & 1 & 0 & -2 \\ -1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 0 & 1 \\ 0 & -2 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}, \tag{3.33}$$

$$\breve{A}^L = \left[\begin{array}{cccc} 0 & 0 & 0 & 1 \end{array}\right] \left[\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 2 & 0 & -2 & 1 \end{array}\right] = \left[\begin{array}{cccc} 2 & 0 & -2 & 1 \end{array}\right].$$

Заметим, что делитель нуля $\breve{A}^R$ содержит ранее указанный столбец (см. пример 3.5).

Умножение полученных матриц по формулам типа (3.25) на произвольные матрицы формирует множества эквивалентных делителей нуля

$$\left\{\breve{A}^L\right\}_{\Pi} = \Pi \left[\begin{array}{cccc} 2 & 0 & -2 & 1 \end{array}\right],$$

$$\left\{\breve{A}^R\right\}_{\Upsilon} = \left[\begin{array}{cc} 0 & -1 \\ 0 & 1 \\ 0 & -2 \\ 1 & 0 \\ 0 & 1 \end{array}\right] \left[\begin{array}{cc} \upsilon_1 & \upsilon_2 \\ \upsilon_3 & \upsilon_4 \end{array}\right].$$

Для формирования матриц преобразования базисов T_y и T_x^{-1} можно использовать два способа, основанные на элементарных преобразованиях матриц.

Напомним, что реализация любого элементарного преобразования над строками (столбцами) матрицы равносильна умножению этой матрицы слева (справа) на некоторую матрицу $P\,(Q)$, называемую *элементарной.*

Матрицы элементарных преобразований получаются из единичной матрицы, если к ней применить соответствующее элементарное преобразование.

Введем символические обозначения для левых и правых элементарных операций и соответствующих матриц:

левые операции:

1) умножение i-й строки матрицы на число c, отличное от нуля: $P' = \{(c)i\}$,

2) прибавление к i-й строке матрицы j-й строки, умноженной на ненулевое число c: $P'' = \{i + (c)j\}$,

3) перестановка местами i-й и j-й строк: $P''' = \{i, j\}$;

правые операции: 1) умножение i-го столбца матрицы на число c, отличное от нуля: $Q' = [(c)i]$, 2) прибавление к i-му столбца матрицы j-го столбца, умноженного на ненулевое число c: $Q'' = [i + (c)j]$, 3) перестановка местами i-го и j-го столбцов: $Q''' = [i, j]$.

Первый способ основан на построении матриц эквивалентных преобразований через элементарные матрицы P_i, Q_j.

Выполнение элементарных преобразований строк и столбцов матрицы A в матричной записи имеет вид

$$P_h \ldots P_1 A Q_1 \ldots Q_v = T_y A T_x^{-1}. \qquad (3.34)$$

Пример 3.8 (Первый способ формирования делителей нуля). Приведем матрицу (3.21) к каноническим базисам, выполняя любую подходящую цепочку элементарных преобразований (при этом используем символические обозначения для элементарных операций)

$$\left[\begin{array}{ccccc} 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 3 & 1 & 0 & 0 \\ 0 & 4 & 2 & 0 & 0 \end{array}\right] \underset{\sim}{[2-1]} \left[\begin{array}{ccccc} 1 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 1 \\ 1 & 2 & 1 & 0 & 0 \\ 0 & 4 & 2 & 0 & 0 \end{array}\right] \underset{\sim}{\{4-(2)3\}}$$

$$\sim \left[\begin{array}{ccccc} 1 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 1 \\ 1 & 2 & 1 & 0 & 0 \\ -2 & 0 & 0 & 0 & 0 \end{array}\right] \underset{\sim}{\{4+(2)1\}} \left[\begin{array}{ccccc} 1 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 1 \\ 1 & 2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] \underset{\sim}{[1-5]}$$

$$\sim \left[\begin{array}{ccccc} 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 1 \\ 1 & 2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] \underset{\sim}{\{3-1\}} \left[\begin{array}{ccccc} 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 1 \\ 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] \underset{\sim}{\{3+(2)2\}}$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \overset{[5+2]}{\sim} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \overset{[5-(2)3]}{\sim}$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \overset{[(-1)2]}{\sim} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Заметим, что последовательность и конкретные варианты использованных элементарных преобразований не влияют на конечный результат.

Сопоставляя каждому элементарному преобразованию соответствующую матрицу элементарных преобразований, запишем выполненные элементарные операции для строк в матричном виде, соблюдая их очередность, и перемножим их. В результате получим матрицу преобразования координат T_y:

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{\{3+(2)2\}} \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{\{3-1\}} \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 2 & 0 & 0 & 1 \end{bmatrix}}_{\{4+(2)1\}} \times$$

$$\times \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -2 & 1 \end{bmatrix}}_{\{4-(2)3\}} = \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 2 & 1 & 0 \\ 2 & 0 & -2 & 1 \end{bmatrix}}_{T_y}.$$

То же самое проделаем с элементарными преобразованиями столбцов и получим матрицу преобразования координат T_x^{-1}:

$$\underbrace{\begin{bmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}}_{[2-1]} \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \end{bmatrix}}_{[1-5]} \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}}_{[5+2]} \times$$

$$\times \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & -2 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}}_{[5-(2)3]} \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}}_{[(-1)2]} =$$

$$= \underbrace{\begin{bmatrix} 1 & 1 & 0 & 0 & -1 \\ 0 & -1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & -2 \\ 0 & 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 0 & 1 \end{bmatrix}}_{T_x^{-1}}.$$

Обе матрицы преобразования координат в силу указанной неединственности отличаются от записанных ранее вариантов (см. пример 3.6). Тем не менее, их использование в формулах (3.31) и (3.32) дает те же значения делителей нуля (3.33).

Поскольку каждая элементарная матрица получается из единичной в результате одноименного элементарного преобразования, то матрицы преобразований T_y, T_x^{-1} можно также построить следующим образом: найдя ряд элементарных преобразований, переводящих A в $\mathcal{A}$, применить все преобразования строк в том же порядке к единичной матрице E_m, а все преобразования столбцов в том же порядке к единичной матрице E_n. Эти преобразования удобно выполнять, используя следующую схему.

Второй способ. Матрица, для которой определяются делители нуля, дополняется двумя единичными матрицами слева и снизу так, что получается конструкция

$$\begin{array}{c|c} E_m & A_{m\times n} \\ \hline & E_n \end{array}\ , \tag{3.35}$$

формально не являющаяся матрицей, а по виду напоминающая развернутый планшет.

Теперь выполняются элементарные преобразования строк и столбцов матрицы A с одновременным преобразованием прилегающих единичных матриц. Цель преобразований заключается в приведении матрицы A к записи в канонических базисах. В результате преобразований получается конструкция

$$\begin{array}{c|c} T_y & \mathcal{A} \\ \hline & T_x^{-1} \end{array}\ , \tag{3.36}$$

где

$$T_y = \begin{bmatrix} \tilde{A}^L_{r\times m} \\ \breve{A}^L_{(m-r)\times m} \end{bmatrix}, \quad T_x^{-1} = \begin{bmatrix} \tilde{A}^R_{n\times r} & \breve{A}^R_{n\times(n-r)} \end{bmatrix}. \tag{3.37}$$

При этом блоки первоначально единичных матриц напротив нулевых блоков матрицы $\mathcal{A}$ будут содержать делители нуля максимального ранга. Левее матрицы $\mathcal{A}$ будет расположен левый делитель, а ниже – правый делитель:

$$\begin{array}{c|c} \begin{bmatrix} \tilde{A}^L_{r\times m} \\ \breve{A}^L_{(m-r)\times m} \end{bmatrix} & \begin{bmatrix} E_r & 0_{r\times(n-r)} \\ 0_{(m-r)\times r} & 0_{(m-r)\times(n-r)} \end{bmatrix} \\ \hline & \begin{bmatrix} \tilde{A}^R_{n\times r} & \breve{A}^R_{n\times(n-r)} \end{bmatrix} \end{array}\ . \tag{3.38}$$

Отсутствие нулевых строк и/или нулевых столбцов у матрицы $\mathcal{A}$, стоящей в правом верхнем углу (3.38), говорит о том, что матрица A не имеет левых и/или правых делителей нуля.

Пример 3.9 (Второй способ формирования делителей нуля (планшет)). Для рассмотренного примера (3.21) конструкция (3.35) примет вид

$$\begin{array}{c|c} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} & \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 3 & 1 & 0 & 0 \\ 0 & 4 & 2 & 0 & 0 \end{bmatrix} \\ \hline & \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \end{array}.$$

Выполнение элементарных преобразований строк и столбцов этого «планшета» приводит к результату

$$\begin{array}{c|c} \breve{A}^L \left\{ \left[\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 2 & 1 & 0 \\ \hline 2 & 0 & -2 & 1 \end{array}\right]\right. & \left[\begin{array}{ccc|cc} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 \end{array}\right] \\ \hline & \left[\begin{array}{ccc} 1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{array} \middle| \underbrace{\begin{array}{cc} 0 & -1 \\ 0 & 1 \\ 0 & -2 \\ 1 & 0 \\ 0 & 1 \end{array}}_{\breve{A}^R} \right] \end{array}. \tag{3.39}$$

Здесь отмечены блоки, равные полученным ранее делителям нуля максимального ранга для матрицы (3.21).

Канонизация матриц

Распишем тождество (3.28) с учетом (3.37) по блокам:

$$\tilde{A}^L A\,\tilde{A}^R = E_r, \quad \breve{A}^L A\,\tilde{A}^R = 0, \quad \tilde{A}^L A\,\breve{A}^R = 0, \quad \breve{A}^L A\,\breve{A}^R = 0.$$

Определение 3.4. Прямоугольные матрицы $\tilde{A}^L$ и $\tilde{A}^R$ ранга r, одновременное умножение на которые слева и справа по формуле

$$\tilde{A}^L_{r\times m} A_{m\times n}\,\tilde{A}^R_{n\times r} = E_r \tag{3.40}$$

приводит прямоугольную матрицу A ранга r к единичной матрице E_r, будем называть соответственно *левым* и *правым канонизаторами*.

Левый и правый канонизаторы характеризуют все линейно независимые комбинации строк и столбцов, соответственно.

При наличии делителей нуля в соответствии со свойством (3.24) каждый из канонизаторов является элементом соответствующего множества канонизаторов, т.е.

$$\left\{\tilde{A}^L\right\}_\Pi = \tilde{A}^L + \Pi\,\breve{A}^L, \qquad \left\{\tilde{A}^R\right\}_\Upsilon = \tilde{A}^R + \breve{A}^R\Upsilon, \tag{3.41}$$

где Π и Υ — матрицы подходящих размеров с произвольными элементами.

Пример 3.10 (Канонизаторы числовой матрицы). Воспользовавшись (3.39) и выполняя соответствующие вычисления, можно убедиться в удовлетворении тождества (3.40) для рассмотренной матрицы (3.21), т.е.

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 2 & 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 3 & 1 & 0 & 0 \\ 0 & 4 & 2 & 0 & 0 \end{bmatrix}\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Определение 3.5. *Сводным канонизатором* (или просто *канонизатором*) прямоугольной матрицы A размера $m\times n$ и ранга r называется матрица размера $n\times m$, равная произведению правого и левого канонизаторов матрицы A:

$$\tilde{A}_{n\times m} = \tilde{A}^R_{n\times r}\tilde{A}^L_{r\times m}. \tag{3.42}$$

Сводный канонизатор характеризует все линейно независимые комбинации строк и столбцов в совокупности.

Пример 3.11 (Сводный канонизатор числовой матрицы). Для матрицы (3.21) из примера 3.5 сводный канонизатор в соответствии с (3.39),(3.42) имеет вид

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 2 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{bmatrix}.$$

Таким образом, любой матрице A можно сопоставить в общем случае неединственную тройку матриц, включающую левый $\breve{A}^L$ и правый $\breve{A}^R$ делители нуля максимального ранга, а также сводный канонизатор $\tilde{A}$, т.е.

$$A \quad\to\quad (\breve{A}^L,\, \tilde{A},\, \breve{A}^R) \tag{3.43}$$

или (при необходимости) четверку матриц

$$A \quad\to\quad (\breve{A}^L,\, \tilde{A}^L,\, \tilde{A}^R,\, \breve{A}^R). \tag{3.44}$$

Такое представление произвольной матрицы будем называть *канонизацией* матрицы.

3.4.1. Односторонние уравнения $AX=B$ и $XA=B$

Применим метод канонизации к решению матричных уравнений.

Теорема 3.4 (Условие разрешимости левостороннего матричного уравнения). *Левостороннее линейное матричное уравнение*

$$AX = B \tag{3.45}$$

с матрицами $A_{m\times n}$, $X_{n\times k}$ и $B_{m\times k}$ разрешимо относительно матрицы X тогда и только тогда, когда левый делитель нуля $\breve{A}^L_{(m-r)\times m}$ матрицы A максимального ранга является и левым делителем нуля матрицы B:

$$\breve{A}^L B = 0. \tag{3.46}$$

Доказательство. Для доказательства воспользуемся теоремой Кронекера – Капели, в соответствии с которой уравнение (3.45) разрешимо тогда и только тогда, когда справедливо тождество $r = \operatorname{rank}\left[\begin{array}{c|c} A & B \end{array}\right]$. Такое возможно тогда и только тогда, когда столбцы матрицы B образованы линейной комбинацией столбцов матрицы A, т.е. существует такая невырожденная матрица S, что

$$B = AS. \tag{3.47}$$

Таким образом, утверждения теоремы Кронекера – Капели и (3.44) эквивалентны

$$\breve{A}^L B = \breve{A}^L AS = 0,$$

что и подтверждает (3.46). □

По аналогии можно доказать еще одну теорему.

Теорема 3.5 (Условие разрешимости правостороннего матричного уравнения). *Правостороннее линейное матричное уравнение*

$$XA = B \tag{3.48}$$

с матрицами $X_{m\times n}$, $A_{n\times k}$ и $B_{m\times k}$ разрешимо относительно матрицы X тогда и только тогда, когда правый делитель нуля $\breve{A}^R_{k\times(k-r)}$ матрицы A максимального ранга является и правым делителем нуля матрицы B:

$$B\breve{A}^R = 0. \tag{3.49}$$

Перейдем теперь к формульному представлению множества решений уравнения (3.45).

Теорема 3.6. *Если левостороннее матричное уравнение (3.45) разрешимо, то все множество его решений определяется формулой*

$$\{\underset{\Upsilon}{\overset{\circ}{X}}\} = \tilde{A}B + \breve{A}^R\Upsilon, \tag{3.50}$$

где Υ — матрица размера $(n-r)\times k$ с произвольными элементами.

В дальнейшем множество (3.50) по названию использованного делителя нуля будем называть *правым множеством решений.* Оно представляет собой сумму двух слагаемых: общего решения $\breve{A}^R\Upsilon$ однородного уравнения

$$AX = 0, \tag{3.51}$$

вытекающего из (3.45) при нулевом значении правой части $B = 0$, и частного решения $\tilde{A}B$ неоднородного уравнения (3.45).

Доказательство. Будем рассматривать общий случай, когда матрица коэффициентов A ранга r имеет оба делителя нуля.

Докажем сначала представление общего решения однородного уравнения. Во-первых, подставляя $\breve{A}^R\Upsilon$ в однородное уравнение (3.51), получим согласно определению правого делителя нуля

$$A\breve{A}^R\Upsilon = 0.$$

Далее, в силу свойств (3.25) правого делителя нуля матрицы A однородное уравнение становится тождеством при любых матрицах Υ подходящего размера. Таким образом, $\breve{A}^R\Upsilon$ действительно является решением однородного уравнения (3.51) при любых матрицах Υ подходящего размера.

Докажем теперь, что любое решение однородного уравнения может быть получено в виде $\breve{A}^R\Upsilon$.

Прежде всего, однородное уравнение (3.51) разрешимо всегда. Введем для этого решения обозначение Ψ. Тогда можно записать тождество

$$A\Psi = 0,$$

которое в соответствии с (3.22), (3.25) имеет место только в том случае, если матрица Ψ представляет собой *какой-либо* правый делитель нуля матрицы A.

В то же время в соответствии со свойством (3.25) любой правый делитель нуля всегда получается умножением делителя нуля максимального ранга $\breve{A}^R$ справа на некоторую матрицу подходящего размера, которую можно обозначить Υ. В результате получаем

$$\Psi = \breve{A}^R\Upsilon.$$

Докажем теперь представление частного решения неоднородного уравнения. Для этого сделаем подстановку

$$A(\tilde{A}B) = B \tag{3.52}$$

и докажем, что это равенство является тождеством.

Умножим обе части равенства (3.52) слева на невырожденную матрицу T_y из (3.26), (3.37). В результате получим эквивалентное равенство

$$\begin{bmatrix} \tilde{A}^L \\ \breve{A}^L \end{bmatrix} A\tilde{A}B = \begin{bmatrix} \tilde{A}^L \\ \breve{A}^L \end{bmatrix} B. \tag{3.53}$$

Раскроем блочную запись в (3.53) и с учетом свойств левого делителя нуля $\breve{A}^L$ получим два равенства

$$\tilde{A}^L A\tilde{A}B = \tilde{A}^L B, \qquad 0 = \breve{A}^L B. \tag{3.54}$$

Второе из полученных равенств повторяет уже доказанное нами утверждение о необходимости и достаточности выполнения условия (3.46) разрешимости уравнения. В первом равенстве раскроем сводный канонизатор по формуле (3.42), в результате чего получим запись

$$\tilde{A}^L A\tilde{A}^R\tilde{A}^L B = \tilde{A}^L B. \tag{3.55}$$

В соответствии с (3.40) произведение первых трех сомножителей в левой части равенства (3.55) тождественно равно единичной матрице E_r. Из этого следует, что (3.55) является тождеством

$$\tilde{A}^L B = \tilde{A}^L B.$$

Поскольку все выполненные преобразования относятся к невырожденным преобразованиям, т.е. эквивалентным, то (3.52) тоже является тождеством, что и требовалось доказать. □

По аналогии можно сформулировать и доказать теорему для правостороннего уравнения.

Теорема 3.7. *Если правостороннее матричное уравнение (3.48) разрешимо, то все множество его решений определяется формулой*

$$\{\overset{\circ}{X}\}_{\Pi} = B\tilde{A} + \Pi\breve{A}^L, \tag{3.56}$$

где Π *– матрица размера* $m \times (n-r)$ *с произвольными элементами.*

Здесь множество решений представлено *левым множеством решений* по названию используемого левого делителя нуля матрицы A.

3.4.2. Двустороннее уравнение $AXC = B$

Будем рассматривать линейное матричное уравнение вида

$$AXC = B, \tag{3.57}$$

где A, B, C — тройка заданных матриц, X — матрица с неизвестными (искомыми) элементами.

Теорема 3.8. *Двустороннее линейное матричное уравнение (3.57) разрешимо относительно матрицы X тогда и только тогда, когда левый делитель нуля матрицы A максимального ранга является и левым делителем нуля матрицы B, а правый делитель нуля матрицы C максимального ранга является и правым делителем нуля матрицы B:*

$$\breve{A}^L B = 0, \qquad B\breve{C}^R = 0. \tag{3.58}$$

Доказательство. По аналогии с односторонними уравнениями (3.45) и (3.48) совместность уравнения (3.57) имеет место тогда и только тогда, когда можно утверждать, что матрица B сформирована, во-первых, из столбцов матрицы A, а во-вторых, из строк матрицы C. Это означает, что по аналогии с (3.47) для матрицы B можно записать следующее разложение на матричные сомножители:

$$B = A\Phi C, \tag{3.59}$$

где Φ – некоторая матрица подходящего размера. Но тогда левый делитель нуля матрицы A и правый делитель нуля матрицы C должны быть соответствующими делителями нуля и для матрицы B. □

Теорема 3.9. *Если двустороннее матричное уравнение (3.57) разрешимо, то все множество его решений определяется формулой*

$$\underset{\Upsilon,\Pi}{\{\overset{\circ}{X}\}} = \tilde{A}B\tilde{C} + \breve{A}^R\Upsilon + \Pi\breve{C}^L, \tag{3.60}$$

где Υ и Π — матрицы подходящих размеров с произвольными элементами.

Здесь слагаемое $\breve{A}^R\Upsilon + \Pi\breve{C}^L$ — общее решение однородного уравнения

$$AXC = 0, \tag{3.61}$$

вытекающего из (3.57) при нулевом значении правой части $B = 0$, слагаемое $\tilde{A}B\tilde{C}$ — частное решение неоднородного уравнения (3.57).

Доказательство. Будем рассматривать общий случай, характеризуемый наличием всех делителей нуля матриц A и C.

Во-первых, непосредственной подстановкой общего решения из (3.60) в (3.61) с использованием свойства (3.24) легко убедится, что

$$\breve{A}^R\Upsilon + \Pi\breve{C}^L$$

действительно является общим решением уравнения (3.61).

Докажем теперь, что любое решение однородного уравнения может быть получено в виде $\breve{A}^R\Upsilon + \Pi\breve{C}^L$.

Для этого проанализируем однородное уравнение (3.61). Левая часть уравнения принимает нулевое значение в четырех случаях:

1. $X = 0$ (тривиальное решение),
2. $X \neq 0, \quad AX = 0,$
3. $X \neq 0, \quad XC = 0,$
4. $X \neq 0, \quad AX \neq 0, \quad XC \neq 0, \quad AXC = 0.$

Покажем, что последний случай обобщает все предыдущие случаи. Указанное тождество может иметь место тогда и только тогда, когда вместе или по отдельности выполняются тождества:

— либо $AX = N\breve{C}^L$,
— либо $\breve{A}^R\Phi = XC$,

где N и Φ – произвольные матрицы подходящих размеров. Рассматривая эти тождества в качестве односторонних матричных уравнений относительно X, в соответствии с (3.50) и (3.56) можно записать

$$\underset{N,\Upsilon}{\{\overset{\circ}{X}\}} = \tilde{A}N\breve{C}^L + \breve{A}^R\Upsilon, \quad \underset{\Phi,\Pi}{\{\overset{\circ}{X}\}} = \breve{A}^R\Phi\,\tilde{C} + \Pi\breve{C}^L. \tag{3.62}$$

Здесь выбор матриц N и Φ в соответствии с (3.46) и (3.49) стеснен условиями

$$\breve{A}^L N\breve{C}^L = 0, \quad \breve{A}^R\Phi\,\breve{C}^R = 0, \tag{3.63}$$

а выбор Υ и Π ничем не ограничен. Полагая, что в общем случае могут быть справедливыми оба решения (3.62), найдем их сумму

$$\underset{\Pi,\Phi,\Upsilon,\Pi}{\{\overset{\circ}{X}\}} = \breve{A}^R(\Upsilon + \Phi\tilde{C}) + (\Pi + \tilde{A}N)\breve{C}^L. \tag{3.64}$$

В силу произвольности матриц Υ и Π ограничивающие условия (3.63) не играют роли в том смысле, что без потери общности можно полагать $N = 0$ и $\Phi = 0$. В результате решение (3.64) приводится к записанному ранее общему решению однородного уравнения. Соответствующим выбором Υ и Π можно получить указанные выше случаи 1 – 3. Представление общего решения однородного уравнения доказано.

Доказательство представления частного решения выполняется аналогично теореме 3.6. □

Алгоритм метода канонизации решения матричных уравнений

Основанный на результатах предыдущего пункта алгоритм решения одностороннего или двустороннего уравнения методом канонизации (3.45), (3.48) или (3.57) сводится к следующему.

Шаг 1. Для матриц коэффициентов A и/или C осуществляется канонизация с использованием «планшета»:

$$A \to (\breve{A}^L, \tilde{A}\,, \breve{A}^R), \qquad C \to (\breve{C}^L, \tilde{C}\,, \breve{C}^R).$$

Шаг 2. В соответствии с теоремами 3.4, 3.5, 3.8 проверяется разрешимость уравнений. Мнемоническое правило сводится к тому, что для матрицы коэффициентов, стоящей левее искомой матрицы, всегда используется левый делитель нуля в качестве левого сомножителя для правой части уравнения. Соответственно для матрицы, стоящей правее – правый делитель нуля как правый сомножитель.

Шаг 3. При удовлетворении условий разрешимости выписывается формула решения (3.50),(3.56) или (3.60), с использованием оставшихся матриц канонизации, т.е. канонизатора и другого делителя нуля.

Отсутствие какого-либо делителя нуля интерпретируется как его нулевое значение.

Пример 3.12 (Решение левостороннего уравнения методом канонизации. Единственное решение). Для иллюстрации алгоритма рассмотрим пример

$$\begin{bmatrix}1&2\\1&1\\2&1\end{bmatrix} X = \begin{bmatrix}5&3\\3&2\\4&3\end{bmatrix}.$$

Осуществление канонизации на шаге 1 алгоритма включает построение и преобразование «планшета»

$$\begin{array}{c|c} \begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix} & \begin{bmatrix}1&2\\1&1\\2&1\end{bmatrix} \\ \hline & \begin{bmatrix}1&0\\0&1\end{bmatrix} \end{array} \to \begin{array}{c|c} \begin{bmatrix}-1&2&0\\1&-1&0\\ \hline 1&-3&1\end{bmatrix} & \begin{bmatrix}1&0\\0&1\\ \hline 0&0\end{bmatrix} \\ \hline & \begin{bmatrix}1&0\\0&1\end{bmatrix} \end{array}.$$

Для канонизаторов получаем

$$\tilde{A}^L = \begin{bmatrix} -1 & 2 & 0 \\ 1 & -1 & 0 \end{bmatrix}, \ \tilde{A}^R = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \ \tilde{A} = \tilde{A}^L = \begin{bmatrix} -1 & 2 & 0 \\ 1 & -1 & 0 \end{bmatrix}.$$

Делители нуля максимального ранга имеют следующий вид:

$$\breve{A}^L = \begin{bmatrix} 1 & -3 & 1 \end{bmatrix}, \quad \breve{A}^R = 0,$$

так как при этом правый делитель содержит нулевое число столбцов (в правой верхней матрице отсутствуют нулевые столбцы).

Выполнение шага 2 подтверждает разрешимость уравнения:

$$\breve{A}^L B = \begin{bmatrix} 1 & -3 & 1 \end{bmatrix} \begin{bmatrix} 5 & 3 \\ 3 & 2 \\ 4 & 3 \end{bmatrix} = \begin{bmatrix} 0 & 0 \end{bmatrix} = 0.$$

На последнем шаге 3 вычисляется решение

$$\overset{\circ}{X} = \begin{bmatrix} -1 & 2 & 0 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} 5 & 3 \\ 3 & 2 \\ 4 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix},$$

оказавшееся единственным в силу равенства нулю делителя нуля $\breve{A}^R$. Ответ легко проверяется подстановкой.

Пример 3.13 (Решение левостороннего матричного уравнения методом канонизации. Множество решений). Решим уравнение

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 2 & 2 \end{bmatrix} X = \begin{bmatrix} 2 & 3 \\ 2 & 3 \\ 4 & 6 \end{bmatrix}. \tag{3.65}$$

Шаг 1:

$$\begin{array}{c|c} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} & \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 2 & 2 \end{bmatrix} \\ \hline & \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \end{array} \sim \begin{array}{c|c} \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix} & \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \\ \hline & \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \end{array}.$$

$$\tilde{A}^L = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}, \quad \tilde{A}^R = \begin{bmatrix} 1 \\ 0 \end{bmatrix},$$

$$\tilde{A} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

$$\breve{A}^L = \begin{bmatrix} -1 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}, \quad \breve{A}^R = \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

Шаг 2:

$$\breve{A}^L B = \begin{bmatrix} -1 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 2 & 3 \\ 4 & 6 \end{bmatrix} = 0.$$

Шаг 3:

$$\{\underset{\Upsilon}{\overset{\circ}{X}}\} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 2 & 3 \\ 4 & 6 \end{bmatrix} +$$

$$+ \begin{bmatrix} -1 \\ 1 \end{bmatrix} \begin{bmatrix} \upsilon_1 & \upsilon_2 \end{bmatrix} = \begin{bmatrix} 2-\upsilon_1 & 3-\upsilon_2 \\ \upsilon_1 & \upsilon_2 \end{bmatrix}.$$

Проверка:

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} 2-\upsilon_1 & 3-\upsilon_2 \\ \upsilon_1 & \upsilon_2 \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 2 & 3 \\ 4 & 6 \end{bmatrix}.$$

Очевидно, что решаемое уравнение становится тождеством при полученном решении и любых υ_i.

Пример 3.14 (Решение двустороннего уравнения методом канонизации). Рассмотрим двустороннее уравнение

$$\begin{bmatrix} 1 & 3 & 1 \\ 0 & 4 & 2 \end{bmatrix} X \begin{bmatrix} 2 & 1 \\ 4 & 2 \end{bmatrix} = \begin{bmatrix} -16 & -8 \\ -20 & -10 \end{bmatrix}. \tag{3.66}$$

Результаты канонизации левой и правой матриц коэффициентов имеют вид

$$\begin{bmatrix} 1 & 3 & 1 \\ 0 & 4 & 2 \end{bmatrix} \to (\breve{A}^L,\, \tilde{A},\, \breve{A}^R) = \left(0,\, \begin{bmatrix} 1 & -0,5 \\ 0 & 0 \\ 0 & 0,5 \end{bmatrix},\, \begin{bmatrix} -1 \\ 1 \\ -2 \end{bmatrix}\right),$$

$$\begin{bmatrix} 2 & 1 \\ 4 & 2 \end{bmatrix} \to (\breve{C}^L,\, \tilde{C},\, \breve{C}^R) = \left(\begin{bmatrix} -2 & 1 \end{bmatrix},\, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix},\, \begin{bmatrix} 1 \\ -2 \end{bmatrix}\right).$$

Правая часть уравнения выбрана таким образом, что удовлетворяются условия разрешимости (3.58). Действительно, выполняются равенства

$$B\breve{C}^R = \begin{bmatrix} -16 & -8 \\ -20 & -10 \end{bmatrix} \begin{bmatrix} 1 \\ -2 \end{bmatrix} = 0,$$

$$\breve{A}^L B = 0 \cdot \begin{bmatrix} -16 & -8 \\ -20 & -10 \end{bmatrix} = 0.$$

Теперь воспользуемся формулой (3.60), в соответствии с которой решение определяется формулой

$$\{\overset{\circ}{X}\}_{\Upsilon,\Pi} = \begin{bmatrix} 1 & -0,5 \\ 0 & 0 \\ 0 & 0,5 \end{bmatrix} \begin{bmatrix} -16 & -8 \\ -20 & -10 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} +$$

$$+ \begin{bmatrix} -1 \\ 1 \\ -2 \end{bmatrix} \begin{bmatrix} \upsilon_1 & \upsilon_2 \end{bmatrix} + \begin{bmatrix} \pi_1 \\ \pi_2 \\ \pi_3 \end{bmatrix} \begin{bmatrix} -2 & 1 \end{bmatrix} =$$

$$= \begin{bmatrix} -3 & 0 \\ 0 & 0 \\ -5 & 0 \end{bmatrix} + \begin{bmatrix} -\upsilon_1 & -\upsilon_2 \\ \upsilon_1 & \upsilon_2 \\ -2\upsilon_1 & -2\upsilon_2 \end{bmatrix} + \begin{bmatrix} -2\pi_1 & \pi_1 \\ -2\pi_2 & \pi_2 \\ -2\pi_3 & \pi_3 \end{bmatrix}.$$

Неограниченный выбор пяти параметров при шести неизвестных элементах дает богатое разнообразие конкретных решений. Так, среди решений присутствуют, например, матрицы

$$\begin{bmatrix} 2 & 1 \\ 15 & 9 \\ 1 & 4 \end{bmatrix},\quad \begin{bmatrix} -5 & -2 \\ 0 & 3 \\ -7 & -5 \end{bmatrix},\quad \begin{bmatrix} 0 & 0 \\ 0 & -1,5 \\ 0 & 0,5 \end{bmatrix} \text{ и т.д.}$$

Проверка решения при произвольных параметрах υ_i и π_j дает

$$\begin{bmatrix} 1 & 3 & 1 \\ 0 & 4 & 2 \end{bmatrix} \begin{bmatrix} -3-\upsilon_1-2\pi_1 & -\upsilon_2+\pi_1 \\ \upsilon_1-2\pi_2 & \upsilon_2+\pi_2 \\ -5-2\upsilon_1-2\pi_3 & -2\upsilon_2+\pi_3 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 4 & 2 \end{bmatrix} =$$

$$= \begin{bmatrix} -16 & -8 \\ -20 & -10 \end{bmatrix}.$$

Задачи и упражнения

1. По системам элементарных делителей матриц A и B найти общее решение уравнения $\tilde{A}X = X\tilde{B}$ ($\tilde{A}$ и $\tilde{B}$ — жордановы нормальные формы матриц A и B соответственно)

   1) $A:\ (\lambda-1)^4,\ (\lambda+1)^3,\ (\lambda-1)^2,\ \lambda+1;$

   $B:\ (\lambda+1)^4,\ (\lambda-1)^2,\ \lambda-1,\ \lambda+1.$

   2) $A:\ (\lambda-1)^4,\ (\lambda-1)^3,\ (\lambda+1)^2,\ \lambda+1;$

   $B:\ (\lambda+1)^4,\ (\lambda-1)^2, \lambda-1, \lambda+1.$

   3) $A:\ (\lambda-3)^4,\ \lambda^3,\ \lambda^2, (\lambda-3)^2,\ \lambda;$

   $B:\ \lambda^5,\ \lambda^3, (\lambda-3)^2,\ (\lambda-3)^2, \lambda.$

   4) $A:\ (\lambda-2)^6,\ (\lambda+1)^4,\ (\lambda-2)^3,\ \lambda+1;$

   $B:\ (\lambda-2)^4,\ (\lambda+1)^3,\ (\lambda-2)^2.$

   5) $A:\ (\lambda-1)^3,\ (\lambda+1)^2,\ \lambda-1,\ \lambda+1;$

   $B:\ (\lambda-1)^5,\ (\lambda-1)^3,\ (\lambda+1)^2,\ \lambda+1.$

2. По системам элементарных делителей матриц A и B (см. задачу 1) подсчитатm количество произвольных параметров, от которых зависит общее решение уравнения $\tilde{A}X = X\tilde{B}$ ($\tilde{A}$ и $\tilde{B}$ — жордановы нормальные формы матриц A и B, соответственно).

3. Решить матричные уравнения $AX = XA$, если

   1) $A = \begin{bmatrix} 4 & -3 & -3 \\ 6 & -5 & -6 \\ 0 & 0 & 1 \end{bmatrix}$, 2) $A = \begin{bmatrix} 13 & -5 & -6 \\ 16 & -7 & -8 \\ 16 & -6 & -7 \end{bmatrix}$,

3) $A = \begin{bmatrix} 2 & -1 & 2 \\ 5 & -3 & 3 \\ -1 & 0 & -2 \end{bmatrix}$, 4) $A = \begin{bmatrix} -1 & 1 & 1 \\ -5 & 21 & 17 \\ 6 & -26 & -21 \end{bmatrix}$,

5) $A = \begin{bmatrix} 3 & 0 & 8 \\ 3 & -1 & 6 \\ -2 & 0 & -5 \end{bmatrix}$, 6) $A = \begin{bmatrix} 1 & -1 & 2 \\ 3 & -3 & 6 \\ 2 & -2 & 4 \end{bmatrix}$.

4. По системе элементарных делителей матрицы A определить, сколько существует линейно -независимых матриц, перестановочных с матрицей A:

1) $(\lambda-1)^7$, $(\lambda+1)^6$, λ^5, $(\lambda-1)^5$, $(\lambda+1)^3$, λ^3, $(\lambda-1)^3$, $(\lambda+1)^2$, $\lambda-1$.

2) $(\lambda-2)^5$, $(\lambda-1)^5$, λ^4, $(\lambda-1)^3$, λ^3, $(\lambda+1)^2$, $(\lambda-2)^2$, λ^2, $\lambda+1$, $\lambda-1$, λ.

3) $(\lambda-1)^6$, $(\lambda+1)^5$, λ^4, $(\lambda-1)^3$, $(\lambda+1)^3$, λ^2, $(\lambda-1)^2$, $\lambda+1$, λ, $\lambda-1$.

4) λ^{10}, $(\lambda-1)^7$, λ^7, $(\lambda+1)^6$, λ^5, $(\lambda-1)^3$, $(\lambda+1)^3$, λ^3, $(\lambda-1)^2$, $\lambda^2, \lambda+1, \lambda$.

5) $(\lambda-1)^9$, $(\lambda+1)^7$, $(\lambda-1)^7$, $(\lambda+1)^5$, $(\lambda-1)^5$, $(\lambda+1)^3$, $(\lambda-1)^3$, $\lambda+1$, $\lambda-1$.

5. Найти все матрицы, перестановочные с матрицей A:

1) $A = \begin{bmatrix} 3 & 2 & -3 \\ 4 & 10 & -12 \\ 3 & 6 & -7 \end{bmatrix}$, 2) $A = \begin{bmatrix} 6 & 6 & -15 \\ 1 & 5 & -5 \\ 1 & 2 & -2 \end{bmatrix}$,

3) $A = \begin{bmatrix} 1 & 1 & -1 \\ -3 & -3 & 3 \\ -2 & -2 & 2 \end{bmatrix}$, 4) $A = \begin{bmatrix} 0 & 2 & 0 & 0 \\ -3 & 5 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & -1 & 3 \end{bmatrix}$,

5) $A = \begin{bmatrix} 4 & -4 & 2 \\ 2 & -2 & 1 \\ -4 & 4 & -2 \end{bmatrix}$, 6) $A = \begin{bmatrix} 4 & 1 & 1 \\ 2 & 4 & 1 \\ 0 & 1 & 4 \end{bmatrix}$,

7) $A = \begin{bmatrix} 9 & -6 & -2 \\ 18 & -12 & -3 \\ 18 & -9 & -6 \end{bmatrix}$, 8) $A = \begin{bmatrix} 3 & 1 & 0 & 0 \\ -1 & 5 & 0 & 0 \\ 0 & 0 & 6 & 2 \\ 0 & 0 & 3 & 7 \end{bmatrix}$,

9) $A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ -1 & 3 & 0 & 0 \\ 0 & 0 & 6 & 2 \\ 0 & 0 & 3 & 7 \end{bmatrix}$, 10) $A = \begin{bmatrix} -1 & 5 & -1 \\ 1 & 0 & 1 \\ 1 & -1 & 3 \end{bmatrix}$,

11) $A = \begin{bmatrix} 0 & 0 & 1 \\ -2 & 1 & 2 \\ 1 & 0 & 2 \end{bmatrix}$.

6. Решить матричные уравнения $AX = XB$ и записать фундаментальную систему решений:

1) $\begin{bmatrix} 4 & -4 & 2 \\ 2 & -2 & 1 \\ -4 & 4 & -2 \end{bmatrix} X = X \begin{bmatrix} 5 & -9 & -4 \\ 6 & -11 & -5 \\ -7 & 13 & 6 \end{bmatrix}$,

2) $\begin{bmatrix} 4 & -4 & 2 \\ 2 & -2 & 1 \\ -4 & 4 & -2 \end{bmatrix} X = X \begin{bmatrix} 3 & -2 \\ -9 & 6 \end{bmatrix}$,

3) $\begin{bmatrix} 2 & -5 & -3 \\ -1 & -2 & -3 \\ 3 & 15 & 12 \end{bmatrix} X = X \begin{bmatrix} 4 & 1 & 1 \\ 2 & 4 & 1 \\ 0 & 1 & 4 \end{bmatrix}$,

4) $\begin{bmatrix} 4 & 1 & 1 \\ -2 & 1 & -2 \\ 1 & 1 & 4 \end{bmatrix} X = X \begin{bmatrix} 2 & -5 & -3 \\ -1 & -2 & -3 \\ 3 & 15 & 12 \end{bmatrix}$,

5) $\begin{bmatrix} 4 & -1 & -2 \\ 2 & 1 & -2 \\ 1 & -1 & 1 \end{bmatrix} X = X \begin{bmatrix} 5 & -3 & 2 \\ 6 & -4 & 4 \\ 4 & -4 & 5 \end{bmatrix}$,

6) $\begin{bmatrix} 0 & 2 \\ -3 & 5 \end{bmatrix} X = X \begin{bmatrix} 4 & 1 & 1 \\ -2 & 1 & -2 \\ 1 & 1 & 4 \end{bmatrix}$,

7) $\begin{bmatrix} 3 & 0 & -1 \\ -2 & 1 & 1 \\ 3 & -1 & -1 \end{bmatrix} X = X \begin{bmatrix} 3 & 8 \\ -2 & -5 \end{bmatrix}$,

8) $\begin{bmatrix} 4 & 1 & 1 \\ 2 & 4 & 1 \\ 0 & 1 & 4 \end{bmatrix} X = X \begin{bmatrix} -4 & 4 & 2 \\ -1 & 1 & 1 \\ -5 & 4 & 3 \end{bmatrix}$.

7. Пусть $AB = BA$, где

$$A = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}, \quad A_1 \in \mathbb{C}^{n\times n}, \quad A_2 \in \mathbb{C}^{m\times m},$$

причем матрицы A_1, A_2 не имеют одинаковых собственных значений. Доказать, что тогда

$$B = \begin{bmatrix} B_1 & 0 \\ 0 & B_2 \end{bmatrix}, \quad B_1 \in \mathbb{C}^{n\times n}, \quad B_2 \in \mathbb{C}^{m\times m}.$$

8. Доказать следующие свойства матричных делителей нуля максимального ранга.

- $(\breve{A}^L)^T = \left(\breve{A^T}\right)^R$, $\quad (\breve{A}^R)^T = \breve{A^T}^L$ (транспонирование)
- Если $\det S \neq 0$, то $\left(\breve{SA}\right)^L = \breve{A}^L S^{-1}$, $\left(\breve{AS}\right)^R = S^{-1}\breve{A}^R$, (умножение матрицы на обратимую матрицу)
- $(\Upsilon\breve{A}^L)A = 0$, $\quad A(\breve{A}^R\Upsilon) = 0$, $\quad \Upsilon = \mathrm{var}$, (умножение делителя нуля на произвольную матрицу)
- $(AK\breve{A}^L)^2 = 0_{n\times n}$, $K = \mathrm{var}$, $(\breve{A}^R NA)^2 = 0_{n\times n}$, $\quad N = \mathrm{var}$, (нильпотентная конструкция матриц индекса 2 с произвольной матрицей)
- если $\operatorname{rank} AB = \operatorname{rank} A$, то $\left(\breve{AB}\right)^L = \breve{A}^L$, если $\operatorname{rank} BA = \operatorname{rank} A$, то $\left(\breve{BA}\right)^R = \breve{A}^R$, (умножение матрицы на другую матрицу с сохранением ранга)
- если $\operatorname{rank} AB < \operatorname{rank} A$, то $\left(\breve{AB}\right)^L \neq \breve{A}^L$, если $\operatorname{rank} BA < \operatorname{rank} A$, то$\left(\breve{BA}\right)^R \neq \breve{A}^R$, (умножение матрицы на другую матрицу с понижением ранга)
- если $m > n$ и $\operatorname{rank} A_{m\times n} = n$, то

$$\det \begin{bmatrix} A_{m\times n} & \left(\breve{A}^L_{(m-n)\times m}\right)^T \end{bmatrix}_{m\times m} \neq 0,$$

  если $m < n$ и $\operatorname{rank} A_{m\times n} = m$, то

$$\det \begin{bmatrix} A_{m\times n} \\ (\breve{A}^R_{n\times(n-m)})^T \end{bmatrix}_{n\times n} \neq 0,$$

  (обратимость сводной матрицы)

- $\left(\breve{0_{m\times n}}\right)^L = K_{m\times m}$, где $\quad K = \mathrm{var}$, $\quad \left(\breve{0_{m\times n}}\right)^R = K_{n\times n}$, где $K = \mathrm{var}$ (делитель нуля для нулевой матрицы)

9. Показать, что в соответствии с изложенной методикой канонизация нулевой матрицы $D = 0_{m\times n}$ произвольного размера $m \times n$ дает следующий результат:

$$\begin{array}{c|c} \underline{E_m} & \dfrac{D_{m\times n}}{E_n} \end{array} \to \begin{array}{c|cc} \tilde{D}^L_{0\times m} & E_0 & 0_{0\times n} \\ \breve{D}^L_{m\times m} = E_m & 0_{m\times 0} & 0_{m\times n} \\ \hline & \tilde{D}^R_{n\times 0} & \breve{D}^R_{n\times n} = E_n \end{array},$$

$$\tilde{D} = \tilde{D}^R_{n\times 0}\tilde{D}^L_{0\times m} = 0_{n\times m}, \quad 0_{m\times n} \to (E_m,\, 0_{n\times m},\, E_n)\,.$$

10. Доказать, что у матрицы-строки нет левых делителей нуля, но она всегда имеет правый делитель нуля вида (3.29), размера $n \times (n-1)$ и ранга $n-1$. Матрица-столбец напротив не имеет правых делителей нуля и всегда имеет левый делитель нуля вида (3.30), размера $(m-1) \times m$ и ранга $m-1$.

11. Доказать, что у квадратной невырожденной матрицы A делители нуля отсутствуют, а односторонние делители единицы A^L и A^R равны друг другу и равны единственной обратной матрице A^{-1}.

12. На примере матрицы

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix}$$

показать, что различные совокупности элементарных преобразований строк соответствующего «планшета» приводят к различным результатам канонизации, но вместе с тем полученные матрицы связаны между собой очевидными линейными соотношениями.

13. Доказать, что при обратимой матрице A алгоритм метода канонизации решения матричных уравнений совпадает с методом Гаусса.

14. Доказать, что все делители нуля произвольного размера и ранга формируются линейным комбинированием строк (столбцов) делителей нуля максимального ранга.

Вопросы для повторения

1. Каким соотношением связаны решение X матричного уравнения $AX = XB$ и решение $\tilde{X}$ матричного уравнения $\tilde{A}\tilde{X} = \tilde{X}\tilde{B}$, если $\tilde{A}$ и $\tilde{B}$ — жордановы нормальные формы матриц A и B, соответственно?
2. Какую структуру имеет решение $\tilde{X}$ матричного уравнения $\tilde{A}\tilde{X} = \tilde{X}\tilde{B}$, где матрицы $\tilde{A}$ и $\tilde{B}$ имеют жорданову нормальную форму?
3. Как выглядит квадратная матрица, имеющая правильную верхнюю треугольную форму? Приведите примеры.
4. Как выглядит прямоугольная матрица, имеющая правильную верхнюю треугольную форму? Приведите пример.
5. Какое количество произвольных параметров в параметрической квадратной $n \times n$- матрице, имеющей правильную верхнюю треугольную форму?
6. Какое количество произвольных параметров в параметрической прямоугольной $m \times n$- матрице, имеющей правильную верхнюю треугольную форму?
7. Чему равно число произвольных параметров в решении $\tilde{X}$ матричного уравнения $\tilde{A}\tilde{X} = \tilde{X}\tilde{B}$, где матрицы $\tilde{A}$ и $\tilde{B}$ имеют жорданову нормальную форму?
8. Чему равно число произвольных параметров в решении решение X матричного уравнения $AX = XB$?
9. В каком случае уравнение $AX = XB$ имеет только нулевое решение?
10. Составьте алгоритм решения матричного уравнения $AX = XB$.
11. Какому матричному уравнению равносильна задача нахождения всех матриц, перестановочных с матрицей A?
12. Как можно найти все матрицы, перестановочные с заданной матрицей?

13. Составьте алгоритм нахождения всех матриц, перестановочных с матрицей A.
14. Чему равно число всех линейно независимых матриц, перестановочных с матрицей $A \in \mathbb{C}^{m\times m}$?
15. Чему равно число всех линейно независимых матриц, перестановочных с матрицей $A \in \mathbb{C}^{m\times m}$, все элементарные делители которой взаимно просты?
16. В каком случае уравнение $AX - XB = C$ имеет единственное решение?
17. Всегда ли имеет решение уравнение $AX = XB$? Уравнение $AX = XA$? Уравнение $AX - XB = C$? Уравнение $AX = B$? Уравнение $XA = B$? Если нет, то при каких условиях уравнение имеет решение? Если да, то при каких условиях это решение единственно?
18. Как найти общее решение уравнения $AX - XB = C$?

19*. Дайте определение левого и правого делителей нуля.

20*. Как для заданной матрицы найти ее делители нуля?

21*. Для всякой ли матрицы существуют делители нуля?

22*. Как определить все делители нуля заданной матрицы?

23*. Дайте определение левого и правого канонизаторов числовой матрицы.

24*. Дайте определение сводного канонизатора числовой матрицы.

25*. Сформулируйте условия разрешимости односторонних, двустороннего матричного уравнений.

26*. Приведите формулы общих решений односторонних, двустороннего матричных уравнений.

27*. Изложите алгоритм метода канонизации решения матричных уравнений.

Необходимо усвоить

Основные понятия и теоремы

- матрица правильной верхней треугольной формы
- теорема о решении уравнения $AX = XB$
- теорема о решении уравнения $AX = XA$
- теорема о количестве линейно-независимых матриц, перестановочных с заданной матрицей
- матричные делители нуля
- канонизатор матрицы
- теорема о решении уравнения $AX = B$
- теорема о решении уравнения $XA = B$
- теорема о решении уравнения $AXB = C$

Основные умения и навыки

- находить решения матричных уравнений $AX = XB$, $AX = XA$, $AX = B$, $XA = B$, $AXB = C$ для заданных матриц A, B и C.

Часть II

ЛИНЕЙНЫЕ ОТОБРАЖЕНИЯ И ИХ МАТРИЦЫ

Глава 4

Сопряженное отображение

Рекомендуемая литература:
[4], [9], [16], [17], [19]

Цель изучения

Освоить понятие и изучить свойства отображения, сопряженного заданному линейному отображению, вид и свойства матриц, соответствующих заданному линейному отображению и его сопряженному.

Следует повторить

- понятие и основные свойства отображений
- матрица линейного отображения
- понятие изоморфизма
- базис векторного пространства
- скалярное произведение
- матрица Грама

4.1. Сопряженное пространство

Пусть V_n — линейное векторное пространство над полем $\mathbb{R}$, $\dim V_n = n$.

Определение 4.1. Числовая функция $y = f(u)$ с аргументами из V_n и значениями из поля $\mathbb{R}$ называется *линейной функцией* на пространстве V_n, если:

$$f(u_1 + u_2) = f(u_1) + f(u_2) \quad \forall\, u_1, u_2 \in V_n;$$
$$f(\alpha \cdot u) = \alpha f(u) \quad \forall\, u \in V_n, \quad \forall\, \alpha \in F. \tag{4.1}$$

Пусть в V_n задан базис

$$E = \{e_1, \ldots, e_n\}.$$

Тогда произвольный вектор $u \in V_n$ может быть разложен по этому базису

$$u = \alpha_1 e_1 + \cdots + \alpha_n e_n, \tag{4.2}$$

а значение линейного функционала f на векторе u согласно (4.1) представляется в виде разложения

$$\begin{aligned} f(u) &= f(\alpha_1 e_1 + \cdots + \alpha_n e_n) = \\ &= \alpha_1 f(e_1) + \cdots + \alpha_n f(e_n) = \\ &= a_1\alpha_1 + \cdots + a_n\alpha_n, \quad a_i = f(e_i). \end{aligned} \tag{4.3}$$

Из (4.3) видно, что при фиксированном базисе E любому набору из n чисел $a_1, \ldots, a_n$ соответствует линейная функция f, причем только одна. Если обозначить через

$$\alpha = \begin{bmatrix} \alpha_1 \\ \ldots \\ \alpha_n \end{bmatrix}$$

координатный столбец вектора u в базисе E, $a = (a_1, \ldots a_n)$ — вектор-строка функции f, то из (4.3) следует представление

$$f(u) = a\alpha. \tag{4.4}$$

Определим на множестве линейных функций операции сложения функций и умножения функции на число:

$$(f+g)(u) = f(u) + g(u), \quad \forall\, u \in V_n;$$
$$(\alpha f)(u) = \alpha f(u), \quad \forall\, u \in V_n, \quad \forall\, \alpha \in F.$$

Если функция f в базисе E определяется числами $(a_1, \ldots, a_n)$, а функция g в базисе E определяется числами $(b_1, \ldots, b_n)$, то функция $f + g$ определяется числами $(a_1 + b_1, \ldots, a_n + b_n)$, а функция αf — числами $(\alpha a_1, \ldots, \alpha a_n)$ в том же базисе E.

Упражнение 4.1. Проверить, что множество линейных функций образует векторное пространство.

Определение 4.2. Пространство линейных функций, определенных на векторном пространстве V_n, называется *сопряженным пространством* к V_n и обозначается V_n^*.

Упражнение 4.2. Доказать, что V_n^* изоморфно n-мерному арифметическому пространству V_n .

Из изоморфности пространств V_n^* и V_n следует, что

$$\dim V_n^* = n, \quad V_n^{**} = V_n.$$

Алгебраические операции сложения и умножения на число в пространствах V_n и V_n^* связаны соотношениями:

$$\begin{aligned} f(u_1 + u_2) &= f(u_1) + f(u_2), & f(\alpha \cdot u) &= \alpha f(u); \\ (f_1 + f_2)(u) &= f_1(u) + f_2(u), & (\alpha \cdot f)(u) &= \alpha f(u). \end{aligned} \tag{4.5}$$

Векторы и функционалы входят в эти соотношения равноправным образом. Тот факт, что в записи $f(u)$ функционал играет роль функции, а вектор — роль аргумента, не имеет существенного значения. Поэтому для записи значения функции f на векторе u иногда применяют обозначение:

$$f(u) = \langle f \,|\, u \rangle. \tag{4.6}$$

С записью (4.6) связана своя терминология. Функции из пространства V_n^* называют *ковекторами*, а само выражение $\langle f \,|\, u \rangle$ называют *скалярным произведением* вектора с ковектором.

Скалярное произведение (4.6) обладает свойством билинейности: оно линейно по первому аргументу и линейно по второму аргументу. Это вытекает из соотношений (4.5), которые теперь записываются так:

$$\begin{aligned} \langle f \,|\, u_1 + u_2 \rangle &= \langle f \,|\, u_1 \rangle + \langle f \,|\, u_2 \rangle, & \langle f \,|\, \alpha \cdot u \rangle &= \alpha \langle f \,|\, u \rangle, \\ \langle f_1 + f_2 \,|\, u \rangle &= \langle f_1 \,|\, u \rangle + \langle f_2 \,|\, u \rangle, & \langle \alpha \cdot f \,|\, u \rangle &= \alpha \langle f \,|\, u \rangle. \end{aligned} \tag{4.7}$$

Свойства (4.7) скалярного произведения (4.6) аналогичны свойствам скалярного произведения векторов в геометрическом пространстве (смотри, например, [7, стр. 290]). Однако, в отличие от такого скалярного произведения, скалярное произведение (4.6) не является симметричным: аргументы в нем принадлежат разным пространствам и их нельзя переставлять. Ковекторы в скалярном произведении (4.6) всегда пишутся слева, а векторы — справа.

4.2. Базис сопряженного пространства

Определение 4.3. Вектор u и ковектор f называются *ортогональными* друг другу, если $\langle f \,|\, u \rangle = 0$.

Пусть в V_n задан базис $E = \{e_1, \ldots, e_n\}$. Рассмотрим в разложении (4.2) вектора $u \in V_n$ по базису E i-ю координату вектора u. В силу однозначности разложения (4.2) при фиксированном выборе базиса в V_n координата α_i — это число, однозначно определяемое вектором u. Рассмотрим отображения $f_i : V_n \mapsto F$, определив их формулой

$$f_i(u) = \alpha_i. \tag{4.8}$$

При сложении векторов их координаты в базисе E складываются, а при умножении вектора на число его координаты умножаются на это число. Отсюда вытекает линейность определенных выше отображений $f_i : V_n \mapsto F$. Значит, с каждым выбором базиса $E = \{e_1, \ldots, e_n\}$ в пространстве V_n связано некоторое семейство n функционалов из V_n^*. Функционалы $f_l, \ldots, f_n$ называются *координатными функционалами* базиса $\{e_1, \ldots, e_n\}$. Для них

$$f_i(e_j) = \langle f_i \,|\, e_j \rangle = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases} \tag{4.9}$$

Соотношения (4.9) называются *соотношениями биортогональности,* а системы векторов $E = \{e_1, \ldots, e_n\}$ и $\mathrm{F} = \{f_1, \ldots, f_n\}$, удовлетворяющие (4.9) — биортогональными, что обозначается как $E \perp \mathrm{F}$.

Упражнение 4.3. Доказать соотношения биортогональности.

Теорема 4.1. *Координатные функционалы $f_1, \ldots, f_n$ линейно независимы и составляют базис в V_n^*.*

Доказательство. Рассмотрим линейную комбинацию координатных функционалов базиса $\{e_1, \ldots, e_n\}$, равную нулю:

$$\alpha_1 \cdot f_1 + \cdots + \alpha_n \cdot f_n = 0. \tag{4.10}$$

Левая часть (4.10) есть нулевой функционал. Значит, его значение на базисном векторе e_j равно нулю:

$$\alpha_1 f_1(e_j) + \cdots + \alpha_n f_n(e_j) = 0. \tag{4.11}$$

В силу соотношений биортогональности (4.9) из n слагаемых в левой части (4.11) остается лишь одно j-е слагаемое, причем $f_j(e_j) = 1$. Поэтому из (4.11) следует $\alpha_j = 0$. Из произвольности j теперь вытекает тривиальность линейной комбинации (4.10) и линейная независимость координатных функционалов $f_1, \ldots, f_n$.

Для завершения доказательства теоремы покажем, что произвольный функционал $f \in V_n^*$ может быть разложен в линейную комбинацию координатных функционалов $f_1, \ldots, f_n$.

Рассмотрим произвольный вектор u из пространства V_n. Тогда из разложения (4.3) для значения функционала f на векторе u с учетом (4.8) получаем

$$\begin{aligned} f(u) &= \alpha_1 f(e_1) + \cdots + \alpha_n f(e_n) = \\ &= f(e_1) f_1(u) + \cdots + f(e_n) f_n(u). \end{aligned} \tag{4.12}$$

Здесь $f(e_1), \ldots, f(e_n)$ — набор чисел из F. В силу произвольности вектора $u \in V_n$ полученное равенство можно переписать как равенство функционалов:

$$f = f(e_1) \cdot f_1 + \cdots + f(e_n) \cdot f_n. \tag{4.13}$$

Формула (4.13) представляет собой разложение произвольного функционала f в линейную комбинацию координатных функционалов f_i. Коэффициенты такой линейной комбинации оказываются равными значениям функционала f на базисных векторах $E = \{e_1, \ldots, e_n\}$. □

Следствие 4.1. *Для любого базиса E пространства V_n существует, и притом только один, базис F пространства V_n^*, такой, что $E \perp \mathrm{F}$.*

Определение 4.4. Базис $\mathrm{F} = \{f_1, \ldots, f_n\}$, построенный из координатных функционалов базиса $\{e_1, \ldots, e_n\}$ в V_n, называется *сопряженным* базисом для E.

Пример 4.1 (Биортогональные базисы). Рассмотрим пространство $V_2 = \mathbb{R}^2$ и его базис

$$E = \{(4,3), (2,1)\}, \quad (4,3) = e_1,\ (2,1) = e_2.$$

Найдем базис

$$\mathrm{F} = \{f_1, f_2\}, \quad \text{такой, что} \quad \mathrm{F} \perp E.$$

Произвольная линейная функция $f \in V_n^*$ действует на вектор $u = (u_1, u_2) \in \mathbb{R}^2$ по правилу:

$$f(u) = f(u_1, u_2) = au_1 + bu_2.$$

Определим неизвестные пока параметры a, b для искомых функций f_1, f_2, пользуясь соотношениями биортогональности (4.9). Для функции f_1 эти условия имеют вид:

$$\begin{cases} \langle f_1, e_1 \rangle = 1, \\ \langle f_1, e_2 \rangle = 0, \end{cases} \Rightarrow \begin{cases} 4a + 3b = 1, \\ \ \ 2a + b = 0. \end{cases}$$

Аналогично для функции f_2 эти условия имеют вид:

$$\begin{cases} \langle f_2, e_1 \rangle = 0, \\ \langle f_2, e_2 \rangle = 1, \end{cases} \Rightarrow \begin{cases} 4a + 3b = 0, \\ \ \ 2a + b = 1. \end{cases}$$

Решим эти системы уравнений:

$$\left[\begin{array}{cc|cc} 4 & 3 & 1 & 0 \\ 2 & 1 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{cc|cc} 0 & 1 & 1 & -2 \\ 2 & 1 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{cc|cc} 1 & 0 & -\frac{1}{2} & \frac{3}{2} \\ 0 & 1 & 1 & -2 \end{array}\right].$$

Записываем искомые функции с найденными параметрами:

$$f_1(u_1, u_2) = -\frac{1}{2}u_1 + u_2, \quad f_2(u_1, u_2) = \frac{3}{2}u_1 - 2u_2.$$

Таким образом, искомый базис $\mathrm{F} = \{-\frac{1}{2}u_1 + u_2, \frac{3}{2}u_1 - 2u_2\}$.

Покажем, что в биортогональных базисах значение скалярного произведения $\langle f \,|\, u \rangle$ вычисляется проще, чем в других базисах.

Пусть $E = \{e_1, \ldots, e_n\}$ и $\mathrm{F} = \{f_1, \ldots, f_n\}$ — два произвольных базиса в V_n и V_n^*, соответственно.

Рассмотрим скалярное произведение вектора $u \in V_n$, разложение которого в базисе E имеет вид (4.2), и ковектора $f \in V_n^*$, разложенного по базису F:

$$f = b_1 f_1 + \cdots + b_n f_n. \tag{4.14}$$

Выразим скалярное произведение $\langle f \,|\, u \rangle$ через координаты ковектора f в базисе F и вектора u в базисе E:

$$\begin{aligned} \langle f \,|\, u \rangle = \langle b_1 f_1 + \cdots + b_n f_n \,|\, \alpha_1 e_1 + \cdots + \alpha_n e_n \rangle &\overset{(4.5)}{=\!=} \\ = \sum_{i,j=1}^{n} \alpha_i b_j \langle f_j \,|\, e_i \rangle. \end{aligned} \tag{4.15}$$

Если базисы F и E биортогональны, т.е. $E \perp \mathrm{F}$, то с учетом соотношений биортогональности (4.9) представление (4.15) упростится и примет вид:

$$\langle f \,|\, u \rangle = \sum_{i=1}^{n} \alpha_i b_i.$$

Таким образом, в биортогональных базисах значение скалярного произведения $\langle f \,|\, u \rangle$ вычисляется намного проще.

Найдем выражения для координат вектора и ковектора в биортогональных базисах. Для этого вычислим

$$\langle f_i \,|\, u \rangle = \langle f_i \,|\, \alpha_1 e_1 + \cdots + \alpha_n e_n \rangle \overset{(4.5),(4.9)}{=\!=\!=\!=} \alpha_i \langle f_i \,|\, e_i \rangle = \alpha_i.$$

Значит,

$$\alpha_i = \langle f_i \,|\, u \rangle,\ i = \overline{1, n}. \tag{4.16}$$

Далее вычислим

$$\langle f \,|\, e_j \rangle = \langle b_1 f_1 + \cdots + b_n f_n \,|\, e_j \rangle \overset{(4.5),(4.9)}{=\!=\!=\!=} b_j \langle f_j \,|\, e_j \rangle = b_j.$$

Значит,

$$b_j = \langle f \,|\, e_j \rangle,\ j = \overline{1, n}. \tag{4.17}$$

Пример 4.2 (Координаты вектора и ковектора). Найдем координаты вектора $u = (1, -1)$ в базисе E из примера 4.1, $u = \alpha_1 e_1 + \alpha_2 e_2$, используя биортогональный базис F, построенный в этом примере. Согласно (4.16) имеем:

$$\alpha_1 = \langle f_1 \,|\, u \rangle = -\frac{1}{2} \cdot 1 + (-1) = -\frac{3}{2},$$
$$\alpha_2 = \langle f_2 \,|\, u \rangle = \frac{3}{2} \cdot 1 - 2 \cdot (-1) = 3\frac{1}{2}.$$

Найдем координаты ковектора $f = 2u_1 - 3u_2$ в базисе F из примера 4.1, $f = b_1 f_1 + b_2 f_2$, используя биортогональный базис E. Согласно (4.17) имеем:

$$b_1 = \langle f \,|\, e_1 \rangle = 2 \cdot 4 - 3 \cdot 3 = -1,$$
$$b_2 = \langle f \,|\, e_2 \rangle = 2 \cdot 2 - 3 \cdot 1 = 1.$$

4.3. Ортогональное дополнение в сопряженном пространстве

Определение 4.5. Пусть V_k — произвольное подпространство векторного пространства V_n. Множество ковекторов из V_n^*, которые ортогональны всем векторам из V_k, называется *ортогональным дополнением* к пространству V_k и обозначается $V_k^\perp$:

$$V_k^\perp = \{ f \in V_n^* : \langle f \,|\, u \rangle = 0, \quad \forall\, u \in V_k \}.$$

Другими словами, ортогональное дополнение к V_k — это множество всех линейных функций из V_n^*, обращающихся в нуль на векторах из V_k.

Заметим, что ортогональное дополнение подпространства V_k лежит в другом пространстве — V_n^*.

Теорема 4.2. *Ортогональное дополнение подпространства V_k является подпространством пространства V_n^*, при этом* $\dim V_k^\perp = n - k$.

Доказательство. Пусть $f \in V_k^\perp$ — произвольный ковектор из ортогонального дополнения. Покажем, что ортогональность ковектора f всем векторам из V_k равносильна ортогональности f векторам любого базиса в V_k.

Пусть $E_k = \{e_1, \ldots, e_k\}$ — некоторый базис в V_k. Тогда для любого вектора $u \in V_k$ справедливо разложение

$$u = \alpha_1 e_1 + \cdots + \alpha_k e_k,$$

и в силу линейности функции f справедливо

$$\langle f \,|\, u \rangle \overset{(4.7)}{=\!=} \alpha_1 \langle f \,|\, e_1 \rangle \ + \cdots + \alpha_k \langle f \,|\, e_k \rangle.$$

Отсюда следует, что

$$\langle f \,|\, u \rangle = 0 \quad \forall\, u \in V_k \Leftrightarrow$$
$$\Leftrightarrow \alpha_1 \langle f \,|\, e_1 \rangle \ + \cdots + \alpha_k \langle f \,|\, e_k \rangle \ = 0, \forall\, \alpha_i,\ i = \overline{1,k} \Leftrightarrow$$
$$\Leftrightarrow \langle f \,|\, e_i \rangle = 0 \quad \forall\, i = \overline{1,k}. \tag{4.18}$$

Таким образом, произвольный ковектор $f \in V_k^\perp$ является решением системы (4.18). Для заданного базиса E_k с базисными векторами вида $e_i = (e_{i1}, e_{i2}, \ldots, e_{in})$, $i = \overline{1,k}$, система (4.18) равносильна системе линейных алгебраических уравнений:

$$\begin{aligned} f_1 e_{11} + \cdots + f_n e_{1n} &= 0, \\ &\ldots \\ f_1 e_{k1} + \cdots + f_n e_{kn} &= 0. \end{aligned} \tag{4.19}$$

Так как векторы $e_1, \ldots e_k$ линейно независимы, то ранг матрицы системы (4.19) равен k. Как известно из линейной алгебры, множество решений системы (4.19) образует линейное пространство размерности $n - k$. Это равносильно доказываемому утверждению. □

Упражнение 4.4. Непосредственно по определению подпространства доказать, что $V_k^\perp$ — подпространство.

Пример 4.3 (Ортогональное дополнение к подпространству). Рассмотрим систему векторов V_2, являющуюся линейной оболочкой, натянутой на векторы $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0) : V_2 = L(e_1, e_2)$. V_2 — подпространство в $\mathbb{R}^3$ с базисом $E = \{e_1, e_2\}$. Покажем, что ортогональным дополнением $V_2^{\perp}$ к V_2 будет множество линейных функций вида

$$f(u) = bu_3. \tag{4.20}$$

Действительно, любой вектор $u = (u_1, u_2, u_3) \in V_2$ представим в виде $u = \alpha_1 e_1 + \alpha_2 e_2$. Вычислим $\langle f \,|\, u\rangle$ для f вида (4.20):

$$\langle f \,|\, u\rangle = \langle f \,|\, \alpha_1 e_1 + \alpha_2 e_2\rangle = \alpha_1 \langle f \,|\, e_1\rangle + \alpha_2 \langle f \,|\, e_2\rangle = \alpha_1 \cdot 0 + \alpha_2 \cdot 0 = 0.$$

Таким образом, любая функция вида (4.20) ортогональна базису подпространства V_2, а значит, и любому вектору этого подпространства, что и требовалось показать.

Пример 4.4 (Ортогональное дополнение и его размерность).
Для заданного подпространства $W = L(e_1, e_2, e_3)$ — линейной оболочки, натянутой на векторы $e_1 = (3, -1, 2, 0, -1)$, $e_2 = (2, 1, -3, 1, 3)$, $e_3 = (7, 1, -4, 2, 5)$ векторного пространства $\mathbb{R}^5$, построим его ортогональное дополнение $W^{\perp}$. Поскольку $e_3 = e_1 + 2e_2$, то в качестве базиса W можно взять векторы e_1 и e_2. Следовательно, $\dim W = 2$, значит $\dim W^{\perp} = 5 - 2 = 3$.

Для нахождения вида функции $f \in W^{\perp}$ воспользуемся равносильностью (4.18).

Пусть $f(x, y, z, t, u) = ax + by + cz + dt + gu$ — произвольная линейная функция, действующая на $\mathbb{R}^5$. Определим, при каких параметрах a, b, c, d, g функция f ортогональна векторам e_1, e_2. Решим систему $\langle f \,|\, e_i\rangle$, $i = 1, 2$:

$$\begin{bmatrix} 3 & -1 & 2 & 0 & -1 \\ 2 & 1 & -3 & 1 & 3 \end{bmatrix} \sim \begin{bmatrix} -3 & 1 & -2 & 0 & 1 \\ 5 & 0 & -1 & 1 & 2 \end{bmatrix},$$

откуда имеем

$$b = 3a + 2c - g, \; d = -a + c - 2g.$$

Таким образом

$$W^{\perp} = \{f \in (\mathbb{R}^5)^* | f(x, y, z, t, u) =$$
$$= ax + (3a + 2c - g)y + cz + (-5a + c - 2g)t + gu, \; a, c, g \in \mathbb{R}\}.$$

В качестве базиса можно выбрать следующие функции:

$$\begin{cases} a = 1, \; c = 0, \; g = 0 : f_1(x, y, z, t, u) = ax + 3y - 5t, \\ a = 0, \; c = 1, \; g = 0 : f_2(x, y, z, t, u) = 2y + z + t, \\ a = 0, \; c = 0, \; g = 1 : f_3(x, y, z, t, u) = -y - 2t + u. \end{cases}$$

4.4. Сопряженное отображение

Ранее мы видели, что заданному векторному пространству V_n соответствует сопряженное ему пространство V_n^*. Заметим, что нет естественного соответствия между заданным вектором $u \in V_n$ и каким-либо конкретным ковектором $f \in V_n^*$: при заданном векторе $u \in V_n$ скалярное произведение $\langle f \,|\, u\rangle$ определено для любого ковектора $f \in V_n^*$, и наоборот.

Рассмотрим два линейных пространства V_n и V_m, а также сопряженные им пространства V_n^* и V_m^*. Будем рассматривать теперь отображения пространств V_n и V_m, а также пространств V_n^* и V_m^* и установим, что между этими отображениями существует взаимно-однозначное соответствие.

Пусть дано некоторое линейное отображение $\mathcal{A} : V_n \to V_m$.

Определение 4.6. Отображение $\mathcal{A}^* : V_m^* \to V_n^*$ называется *сопряженным отображением* для $\mathcal{A}$, если для любого $u \in V_n$ и для любого $g \in V_m^*$ выполняется соотношение:

$$\langle \mathcal{A}^*(g) \,|\, u\rangle = \langle g \,|\, \mathcal{A}(u)\rangle. \tag{4.21}$$

Заметим, что сопряженное отображение действует, можно так сказать, в направлении, обратном к первоначальному отображению.

Теорема 4.3. *Для любого заданного линейного отображения $\mathcal{A}$ сопряженное отображение $\mathcal{A}^*$ существует, линейно и единственно.*

Доказательство. Построим отображение $\mathcal{A}^*$, удовлетворяющее (4.21) для заданного отображения $\mathcal{A}$. Для того чтобы задать отображение $\mathcal{A}^* : V_m^* \to V_n^*$, надо для каждого функционала $g \in V_m^*$ указать соответствующий ему функционал

$$f = \mathcal{A}^*(g) \in V_n^*, \tag{4.22}$$

являющийся образом g при отображении $\mathcal{A}^*$. Но задать функционал из V_n^* — это значит определить его действие на произвольный вектор $u \in V_n$. Зададим правило этого действия. С учетом определяющего соотношения (4.21) для сопряженного отображения имеем:

$$f(u) = \langle f \,|\, u\rangle \overset{(4.22)}{=\!=} \langle \mathcal{A}^*(g) \,|\, u\rangle \overset{(4.21)}{=\!=} \langle g \,|\, \mathcal{A}(u)\rangle. \tag{4.23}$$

Из (4.23) следует, что для заданного функционала $g \in V_m^*$ действие искомого функционала $f \in V_n^*$ на произвольный вектор $u \in V_n$ определяется следующим образом. Сначала применяем отображение $\mathcal{A} : V_n \to V_m$ к вектору $u \in V_n$, в результате чего получаем его образ — вектор $\mathcal{A}(u) \in V_m$. Затем вычисляем значение заданного функционала $g \in V_m^*$ на векторе $\mathcal{A}(u) \in V_m$ и получаем значение $\langle g \,|\, \mathcal{A}(u)\rangle$, что и является результатом действия f на u.

Таким образом, сопряженное отображение определяется как композиция

$$\mathcal{A}^*(g) = g \circ \mathcal{A}.$$

Проверим, что определенное таким образом отображение является линейным. Для этого проверим следующие условия:

1) $\mathcal{A}^*(g_1 + g_2) = \mathcal{A}^*(g_1) + \mathcal{A}^*(g_2)$;
2) $\mathcal{A}^*(\alpha g) = \alpha \mathcal{A}^*(g)$.

Докажем 1):

$$\langle \mathcal{A}^*(g_1 + g_2) \,|\, u\rangle = \langle g_1 + g_2 \,|\, \mathcal{A}(u)\rangle = \langle g_1 \,|\, \mathcal{A}(u)\rangle + \langle g_2 \,|\, \mathcal{A}(u)\rangle =$$
$$= \langle \mathcal{A}^*(g_1) \,|\, u\rangle + \langle \mathcal{A}^*(g_2) \,|\, u\rangle = \langle \mathcal{A}^*(g_1) + \mathcal{A}^*(g_2) \,|\, u\rangle, \ \forall\, u \in V_n,$$

что и доказывает 1).

Докажем 2):

$$\langle \mathcal{A}^*(\alpha g) \,|\, u\rangle = \langle \alpha g \,|\, \mathcal{A}(u)\rangle = \alpha\langle g \,|\, \mathcal{A}(u)\rangle =$$
$$= \alpha\langle \mathcal{A}^*(g) \,|\, u\rangle = \langle \alpha\mathcal{A}^*(g) \,|\, u\rangle, \quad \forall\, u \in V_n, \forall\, \alpha \in \mathbb{R},$$

что доказывает 2).

Докажем единственность $\mathcal{A}^*$.

Пусть $\mathcal{A}_1^*$, $\mathcal{A}_2^*$ — два отображения, для которых справедливо (4.21). Тогда $\langle \mathcal{A}_1^*(g) \,|\, u\rangle = \langle g \,|\, \mathcal{A}(u)\rangle = \langle \mathcal{A}_2^*(g) \,|\, u\rangle$ для всех $g \in V_n^*$, $u \in V_n$. Отсюда следует, что $\langle (\mathcal{A}_1^* - \mathcal{A}_2^*)(g) \,|\, u\rangle = 0$. Фиксируем g и будем менять u. Тогда элемент $(\mathcal{A}_1^* - \mathcal{A}_2^*)(g) \in V_n^*$ как линейная функция на V_n принимает только нулевые значения, и, значит, равен нулю. Поэтому $\mathcal{A}_1^* = \mathcal{A}_2^*$, что доказывает единственность и завершает доказательство теоремы. □

Пример 4.5 (Сопряженное отображение). Для отображения $\mathcal{A} : V \to W$, $\quad V = \mathbb{R}^3$, $W = \mathbb{R}^5$, такого, что

$$\mathcal{A}(x_1, x_2, x_3) = (x_1 + x_2,\ x_2 + x_3,\ x_1 + x_3,\ x_1, -x_3)$$

и ковектора $g \in W^*$:

$$g(y_1, y_2, y_3, y_4, y_5) = 2y_1 - y_2 - y_3 - y_4 + y_5$$

найдем его образ $f = \mathcal{A}^*(g) \in V^*$ при сопряженном отображении $\mathcal{A}^*(g)$:

$$f(x_1, x_2, x_3) \overset{(4.23)}{=\!=} \langle \mathcal{A}^*(g) \,|\, (x_1, x_2, x_3)\rangle = \langle g \,|\, \mathcal{A}(x_1, x_2, x_3)\rangle =$$
$$= \langle 2y_1 - y_2 - y_3 - y_4 + y_5 \,|\, x_1 + x_2, x_2 + x_3, x_1 + x_3, x_1 - x_3\rangle =$$
$$= 2(x_1 + x_2) - (x_2 + x_3) - (x_1 + x_3) - x_1 + (-x_3) = x_2 - 3x_3.$$

Пусть в пространствах V_n и V_m выбраны базисы $E = \{e_1, \ldots, e_n\}$ и $H = \{h_1, \ldots, h_m\}$. Этим базисам однозначно соответствуют биортогональные базисы $E^\perp = \{e_1^\perp, \ldots, e_n^\perp\}$ и $H^\perp = \{h_1^\perp, \ldots, h_m^\perp\}$ сопряженных пространств V_n^* и V_m^*.

Пусть задано линейное отображение $\mathcal{A} : V_n \to V_m$ и $\mathcal{A}^* : V_m^* \to V_n^*$ — сопряженное ему отображение.

С каждой парой базисов E, H линейных пространств V_n и V_m и линейным отображением $\mathcal{A} : V_n \to V_m$ связана матрица этого отображения. Напомним, что *матрицей линейного отображения* в заданных базисах называется матрица $A = (a_{ij})_{m\times n}$, j-й столбец которой состоит из кординат вектора $\mathcal{A}(e_j)$, т.е. кординат образа j-го базисного вектора:

$$\mathcal{A}(e_j) = a_{1j}h_1 + \cdots + a_{mj}h_m = \sum_{i=1}^{m} a_{ij}h_i$$

в базисе H.

Изучим связь между матрицами исходного и сопряженного отображения.

Пусть отображение $\mathcal{A}$ в базисах E, H имеет матрицу $A = (a_{ij})_{m\times n}$. Определим структуру матрицы $\Phi = (\phi_{ij})_{n\times m}$ сопряженного отображения $\mathcal{A}^*$ в биортогональных базисах $E^\perp, H^\perp$.

Теорема 4.4. *Пусть $\mathcal{A} : V_n \to V_m$ — линейное отображение, E и H — базисы пространств V_n и V_m, соответственно, $E^\perp$ и $H^\perp$ — биортогональные базисы пространств V_n^* и V_m^*. Тогда если отображение $\mathcal{A}$ в базисах E и H имеет матрицу A, то сопряженное отображение $\mathcal{A}^*$ в биортогональных базисах V_n^* и V_m^* имеет матрицу A^T.*

Доказательство. Согласно определению матрицы отображений $\mathcal{A}$ и $\mathcal{A}^*$ определяются из разложений

$$\mathcal{A}(e_j) = \sum_{i=1}^{m} a_{ij}h_i,$$
$$\mathcal{A}^*(h_i^\perp) = \sum_{q=1}^{n} \phi_{qi}e_q^\perp. \tag{4.24}$$

Из определяющего соотношения сопряженного отображения (4.21) имеем:

$$\langle \mathcal{A}^*(h_i^\perp) \,|\, e_j\rangle = \langle h_i^\perp \,|\, \mathcal{A}(e_j)\rangle. \tag{4.25}$$

Вычислим отдельно левую и правую части этого равенства, используя (4.24):

$$\langle \mathcal{A}^*(h_i^\perp) \,|\, e_j\rangle \overset{(4.24)}{=\!=} \langle \sum_{q=1}^{n} \phi_{qi}e_q^\perp \,|\, e_j\rangle = \sum_{q=1}^{n} \phi_{qi}\langle e_q^\perp \,|\, e_j\rangle = \phi_{ji},$$

$$\langle h_i^\perp \,|\, \mathcal{A}(e_j)\rangle \overset{(4.24)}{=\!=} \langle h_i^\perp \,|\, \sum_{i=1}^{m} a_{ij}h_i)\rangle = \sum_{k=1}^{m} a_{kj}\langle h_i^\perp \,|\, h_k)\rangle = a_{ij}.$$

Подставив полученные выражения в (4.25), получим

$$\phi_{ji} = a_{ij},$$

что означает, что $\Phi = A^T$. $\square$

Пример 4.6 (Матрица сопряженного отображения). Рассмотрим два линейных пространства $V_2 = \mathbb{R}^2$ и $V_3 = \mathbb{R}^3$ с базисами

$$E = \{e_1,\, e_2\},\ e_1 = (2, 1), e_2 = (3, -1), \tag{4.26}$$

$$H = \{h_1, h_2, h_3\},\ h_1 = (1, 2, 0), h_2 = (-1, 0, 1), h_3 = (0, 0, -1)$$

и линейное отображение $\mathcal{A} : V_2 \to V_3 :$

$$\mathcal{A}(x_1, x_2) = (x_1 + x_2,\, -x_1,\, -x_2).$$

Найдем матрицу сопряженного отображения $\mathcal{A}^* : V_3^* \to V_2^*$ в биортогональных базисах $E^\perp$, $H^\perp$.

Биортогональный базис $E^\perp = \{e_1^\perp, e_2^\perp\}$ найдем аналогично примеру 4.1. Для нахождения $E^\perp$ решаем систему:

$$\left[\begin{array}{rr|rr} 2 & 1 & 1 & 0 \\ 3 & -1 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{rr|rr} 5 & 0 & 1 & 1 \\ 3 & -1 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{rr|rr} 1 & 0 & \frac{1}{5} & \frac{1}{5} \\ 0 & 1 & \frac{3}{5} & -\frac{2}{5} \end{array}\right].$$

Отсюда следует, что $E^\perp$ состоит из ковекторов:

$$\begin{aligned} e_1^\perp(x_1, x_2) &= \frac{1}{5}x_1 + \frac{3}{5}x_2 = \frac{1}{5}(x_1 + 3x_2), \\ e_2^\perp(x_1, x_2) &= \frac{1}{5}x_1 - \frac{2}{5}x_2 = \frac{1}{5}(x_1 - 2x_2). \end{aligned} \tag{4.27}$$

Для нахождения $H^\perp$ решаем систему:

$$\left[\begin{array}{rrr|rrr} 1 & 2 & 0 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{rrr|rrr} 0 & 2 & 0 & 1 & 1 & 1 \\ -1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & -1 \end{array}\right] \sim$$

$$\sim \left[\begin{array}{rrr|rrr} 1 & 0 & 0 & 0 & -1 & -1 \\ 0 & 1 & 0 & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 1 & 0 & 0 & -1 \end{array}\right].$$

Отсюда следует, что $H^\perp = \{h_1^\perp,\ h_2^\perp,\ h_3^\perp\}$ состоит из ковекторов:

$$\begin{aligned} h_1^\perp(x_1, x_2, x_3) &= \frac{1}{2}x_2, \\ h_2^\perp(x_1, x_2, x_3) &= -x_1 + \frac{1}{2}x_2, \\ h_3^\perp(x_1, x_2, x_3) &= -x_1 + \frac{1}{2}x_2 - x_3. \end{aligned} \tag{4.28}$$

Найдем матрицу A_E^H отображения $\mathcal{A}$ в базисах E и H. Для этого вычислим образы базисных векторов базиса E при отображении $\mathcal{A}$:

$$\mathcal{A}(e_1) = (3, -2, -1);$$
$$\mathcal{A}(e_2) = (2, -3, 1).$$

Обозначим $\alpha^i = \{\alpha_1^i, \alpha_2^i, \alpha_3^i\}^T$ — искомые координатные столбцы вектора $\mathcal{A}(e_i)$, $i = 1, 2$ в базисе H:

$$H \cdot \alpha^i = \mathcal{A}(e_i),\ i = 1, 2. \tag{4.29}$$

Найдем коодинаты α_j^i, $i = 1, 2$, $j = \overline{1, 3}$. Для этого решаем (4.29) относительно α^i, $i = 1, 2$:

$$\left[\begin{array}{rrr|rr} 1 & -1 & 0 & 3 & 2 \\ 2 & 0 & 0 & -2 & -3 \\ 0 & 1 & -1 & -1 & 1 \end{array}\right] \sim \left[\begin{array}{rrr|rr} 0 & -1 & 0 & 4 & \frac{7}{2} \\ 1 & 0 & 0 & -1 & -\frac{3}{2} \\ 0 & 1 & -1 & -1 & 1 \end{array}\right] \sim$$

$$\sim \left[\begin{array}{rrr|rr} 1 & 0 & 0 & -1 & -\frac{3}{2} \\ 0 & 1 & 0 & -4 & -\frac{7}{2} \\ 0 & 0 & 1 & -3 & -\frac{9}{2} \end{array}\right].$$

Таким образом, матрица отображения $\mathcal{A}$ в базисах E и H имеет вид:

$$A_E^H = \begin{bmatrix} -1 & -\frac{3}{2} \\ -4 & -\frac{7}{2} \\ -3 & -\frac{9}{2} \end{bmatrix}.$$

Тогда матрица $A_{E^\perp}^{H^\perp}$ сопряженного отображения $\mathcal{A}^*$ в базисах $E^\perp$ и $H^\perp$ имеет вид:

$$A_{E^\perp}^{H^\perp} = -\frac{1}{2}\begin{bmatrix} 2 & 8 & 6 \\ 3 & 7 & 9 \end{bmatrix}.$$

Свойства сопряженных отображений [4, стр. 11]

Переход к сопряженному отображению связан следующим образом с алгебраическими операциями над отображениями.

Теорема 4.5.

Если определены отображения $\mathcal{B}\mathcal{A}$, $\mathcal{A}^{-1}$, или $\mathcal{A} + \mathcal{B}$, то справедливо

$$\mathcal{B} \circ \mathcal{A}^* = \mathcal{A}^* \circ \mathcal{B}^*, \tag{4.30}$$

$$(\mathcal{A}^{-1})^* = (\mathcal{A}^*)^{-1}, \tag{4.31}$$

$$(\mathcal{A}+\mathcal{B})^* = \mathcal{A}^* + \mathcal{B}^*, \tag{4.32}$$

$$(\alpha\mathcal{A})^* = \alpha\mathcal{A}^*. \tag{4.33}$$

Доказательство. Для доказательства формул используем координатную запись отображений.

Пусть определены отображения $\mathcal{A} : V \mapsto W$, $\mathcal{B} : W \mapsto U$, и в линейных пространствах V, W, U выбраны базисы E, L, H, соответственно. Тогда отображение $\mathcal{B} \circ \mathcal{A}$ в паре базисов E, H имеет матрицу BA, где A — матрица $\mathcal{A}$ в базисах E, L, а B — матрица $\mathcal{B}$ в базисах L, H. Поскольку $(BA)^T = A^T B^T$ для любых матриц A и B, для которых определено произведение BA, то отображение $\mathcal{B} \circ \mathcal{A}^*$ в базисах, биортогональных H, E, имеет матрицу $A^T B^T$, и, следовательно, равно произведению $\mathcal{A}^* \circ \mathcal{B}^*$. Таким образом, формула (4.30) доказана.

Остальные формулы доказываются аналогично. □

Упражнение 4.5. Доказать формулы (4.31)-(4.33).

4.5. Сопряженное отображение евклидовых пространств

Пусть $\mathcal{E}_n$ и $\mathcal{E}_m$ — евклидовы пространства. Рассмотрим линейное отображение $\mathcal{A} : \mathcal{E}_n \to \mathcal{E}_m$ и сопряженное отображение $\mathcal{A}^*$.

Наиболее существенная отличительная черта евклидова пространства состоит в том, что его можно отождествить с сопряженным ему пространством. Такое отождествление возможно благодаря тому, что существует естественный, не зависящий от выбора базиса, изоморфизм пространств $\mathcal{E}_n^*$ и $\mathcal{E}_n$.

Теорема 4.6. *Евклидово пространство изоморфно своему сопряженному пространству. То есть существует изоморфизм $is : \mathcal{E}_n^* \to \mathcal{E}_n$, который каждой функции $f \in \mathcal{E}_n^*$ ставит в соответствие вектор $v \in \mathcal{E}_n$, при этом*

$$\langle f \,|\, u\rangle = \langle v, u\rangle, \quad \forall\, u \in \mathcal{E}_n. \tag{4.34}$$

Доказательство. Пусть $E = \{e_1, \ldots, e_n\}$ — заданный базис пространства $\mathcal{E}_n$. Для произвольного вектора $u \in E_n$ его разложение по базису E:

$$u = \beta_1 e_1 + \cdots + \beta_n e_n \tag{4.35}$$

определяет координатный столбец вектора u в базисе E:

$$\beta = \begin{bmatrix} \beta_1 \\ \ldots \\ \beta_n \end{bmatrix}.$$

Рассмотрим линейную функцию $f \in \mathcal{E}_n^*$, и пусть a — вектор-строка этой функции в базисе E такая, что согласно (4.4) $f(u) = a\beta$.

Обозначим

$$\alpha = \begin{bmatrix} \alpha_1 \\ \ldots \\ \alpha_n \end{bmatrix}$$

— координатный столбец вектора v в базисе E.

Запишем скалярное произведение $\langle v, u\rangle$ в координатной форме:

$$\langle v, u\rangle = \alpha^T G_E \beta, \tag{4.36}$$

где G_E — матрица Грама базиса E.

Далее, согласно (4.4) $\langle a, u\rangle = f(u) = a\beta$. Запишем (4.34) в координатной форме:

$$\langle f \,|\, u\rangle = \langle a, u\rangle \overset{(4.36)}{\Leftrightarrow} a\beta = \alpha^T G_E \beta \Leftrightarrow a = \alpha^T G_E.$$

Поскольку матрица Грама G_E базиса E невырожденная, то последнее равенство можно решить относительно α. В силу симметричности G_E имеем

$$\alpha = (G_E^T)^{-1} a^T = G_E^{-1} a^T. \tag{4.37}$$

Здесь a^T — координатный столбец функции f в базисе $E^\perp$ сопряженного пространства $\mathcal{E}_n^*$, ортогональном базису E. Итак, из (4.37) следует, что пространства $\mathcal{E}_n$ и $\mathcal{E}_n^*$ изоморфны. □

Таким образом, евклидово пространство изоморфно своему сопряженному. Поэтому можно отождествить евклидово пространство $\mathcal{E}_n$ с сопряженным ему пространством $\mathcal{E}_n^*$.

Определение 4.7. Отображение $\mathcal{A}^* : \mathcal{E}_m \to \mathcal{E}_n$, определяемое равенством

$$\langle \mathcal{A}^*(v), u \rangle = \langle v, \mathcal{A}u \rangle, \ \forall\, u \in \mathcal{E}_n, \forall\, v \in \mathcal{E}_m, \tag{4.38}$$

называется *сопряженным* отображению $\mathcal{A} : \mathcal{E}_n \to \mathcal{E}_m$.

Заметим, что в левой части равенства стоит скалярное произведение в пространстве $\mathcal{E}_n$, а в правой — в пространстве $\mathcal{E}_m$.

В теореме 4.4 была установлена связь между матрицами отображения $\mathcal{A} : V_n \to V_m$ и сопряженного отображения $\mathcal{A}^* : V_m^* \to V_n^*$ в биортогональных базисах произвольных линейных пространств V_n и V_m. Конкретизируем этот результат для евклидовых пространств.

Отождествив $\mathcal{E}_n$ и $\mathcal{E}_n^*$, $\mathcal{E}_m$ и $\mathcal{E}_m^*$, мы выбираем в отождествленных пространствах один и тот же базис, а не два биортогональных друг другу базиса. Таким образом, чтобы для евклидовых пространств получить результат, аналогичный теореме 4.4, надо потребовать, чтобы базис в $\mathcal{E}_n$ совпадал со своим биортогональным, т.е. был ортонормированным.

Таким образом справедлива

Теорема 4.7. *Если отображение $\mathcal{A} : \mathcal{E}_n \to \mathcal{E}_m$ в ортонормированных базисах имеет матрицу A, то его сопряженное отображение $\mathcal{A}^*$ в тех же базисах имеет матрицу A^T.*

Пример 4.7 (Матрица сопряженного отображения евклидовых пространств). Рассмотрим два евклидовых пространства $\mathcal{E}_2$ и $\mathcal{E}_3$ и их базисы $E_2 = \{e_1^2, e_2^2\}$ и $E_3 = \{e_1^3, e_2^3, e_3^3\}$,

$H_2 = \{h_1^2, h_2^2\}$ и $H_3 = \{h_1^3, h_2^3, h_3^3\}$, причем E_2 и E_3 — ортонормированные, а базисы H_2 и H_3 связаны с E_2 и E_3 соотношениями:

$$\begin{aligned} e_1^2 &= h_1^2 - h_2^2,\\ e_2^2 &= h_1^2 + h_2^2,\\ e_1^3 &= h_1^3 - h_2^3,\\ e_2^3 &= -2h_2^3 - h_3^3,\\ e_3^3 &= h_1^3 - h_3^3. \end{aligned}$$

Отсюда следует, что матрица S_2 перехода от базиса H_2 к базису E_2 имеет вид

$$S_2 = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}, \quad S_2^{-1} = \frac{1}{2}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix},$$

а матрица S_3 перехода от базиса H_3 к базису E_3 имеет вид

$$S_3 = \begin{bmatrix} 1 & 0 & 1 \\ -1 & -2 & 0 \\ 0 & -1 & -1 \end{bmatrix}, \quad S_3^{-1} = \frac{1}{3}\begin{bmatrix} 2 & -1 & 2 \\ -1 & -1 & -1 \\ 1 & 1 & -2 \end{bmatrix}.$$

Пусть линейное отображение $\mathcal{A} : \mathcal{E}_2 \to \mathcal{E}_3$, в базисах H_2 и H_3 имеет вид

$$A_H = \begin{bmatrix} 1 & 1 \\ -1 & 1 \\ 1 & 0 \end{bmatrix}.$$

Найдем матрицу сопряженного отображения $\mathcal{A}^* : \mathcal{E}_3^* \to \mathcal{E}_2^*$ в базисах H_2 и H_3.

Сначала найдем матрицу отображения $\mathcal{A}$ в базисах E_2 и E_3:

$$A_E = S_3^{-1} A_H S_2 =$$

$$= \frac{1}{3}\begin{bmatrix} 2 & -1 & 2 \\ -1 & -1 & -1 \\ 1 & 1 & -2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 4 & 6 \\ 1 & -3 \\ -4 & -2 \end{bmatrix}.$$

Так как базисы E_2 и E_3 ортонормированные, то согласно теореме 4.7 матрица A_E^* сопряженного отображения в этих базисах имеет вид

$$A_E^* = (A_E)^T = \frac{1}{3}\begin{bmatrix} 4 & 1 & -4 \\ 6 & -3 & -2 \end{bmatrix}.$$

Тогда матрица A_H^* сопряженного отображения в базисах H_2 и H_3 имеет вид:

$$A_H^* = S_2^{-1} A_E^* S_3 = \frac{1}{2}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \times \frac{1}{3}\begin{bmatrix} 4 & 1 & -4 \\ 6 & -3 & -2 \end{bmatrix} \times$$
$$\times \begin{bmatrix} 1 & 0 & 1 \\ -1 & -2 & 0 \\ 0 & -1 & -1 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} -3 & -3 & 0 \\ 6 & 5 & 8 \end{bmatrix}.$$

Задачи и упражнения

1. Заданы базис $E = (e_1, \ldots, e_n)$ векторного пространства V, вектор $u \in V$ и линейный функционал $f \in V^*$. Найти биортогональный базис $E^\perp$, координаты функции f в нем и координаты вектора u в базисе E.

   1) $V = \mathbb{R}^3$, $e_1 = (2,-1,0)$, $de_2 = (1,3,-2)$, $e_3 = (1,-1,2)$;
   $u = (2,-1,5)$, $f(x,y,z) = 3x - 5y + z$;
   2) $V = \mathbb{R}^3$, $e_1 = (3,0,-3)$, $e_2 = (1,3,-2)$, $e_3 = (-5,2,0)$;
   $u = (7,-2,0)$, $f(x,y,z) = x - 3z$;
   3) $V = \mathbb{R}^4$, $e_1 = (1,0,-1,0)$, $e_2 = (0,1,0,-1)$, $e_3 = (1,1,1,1)$,
   $e_4 = (1,-1,1,-1)$; $u = (1,2,3,4)$, $f(x,y,z) = 2x - y + z - 3t$.

2. Для заданного подпространства W векторного пространства

   $\mathbb{R}^5$ найти его ортогональное дополнение $W^\perp$, указать его базис и размерность:

   1) $W = \{(2x-3y+z,\ x-2y,\ z,\ x+y+z,\ x-5z),\ x,y,z \in \mathbb{R}\}$;
   2) $W = \{(x-2y+3z,\ 2x-y+z,\ 3x-3y+4z,\ x+y-2z,$
   $5x+4y+5z),\ x,y,z \in \mathbb{R}\}$;
   3) $W = L(e_1,e_2,e_3)$, $e_1 = (-2,3,1,0,2)$,
   $e_2 = (3,-1,0,1,3)$, $e_3 = (1,0,2,5,-7)$.

3. Для заданного линейного отображения $\mathcal{A} : V \to W$, и заданного ковектора $g \in W^*$ найти его образ $f = \mathcal{A}^*(g) \in V^*$ при сопряженном отображении $\mathcal{A}^*(g)$.

   1) $V = \mathbb{R}^3$, $W = \mathbb{R}^2$, $\mathcal{A}(x,y,z) = (x-z, z-y)$, $g(x,y) = x+y$;
   2) $V = \mathbb{R}^2$, $W = \mathbb{R}^3$, $\mathcal{A}(x,y) = (2x-y, x, x+y)$,
   $g(x,y,z) = 2x - y - 2z$;
   3) $V = \mathbb{R}^5$, $W = \mathbb{R}^3$, $\mathcal{A}(x,y,z,t,w) = (x-2y+w, y-2t+z,$
   $x+y+z+t+w)$, $g(x,y,z) = 3x - 5y + 6z$.

4. Линейное отображение $\mathcal{A}$ действует из векторного пространства V в векторное пространство W, E — базис пространства V, H — базис пространства W. Найти биортогональные базисы для базисов E и H и матрицу сопряженного отображения $\mathcal{A}^*$ в биортогональных базисах.

   1) $V = \mathbb{R}^2$, $W = \mathbb{R}^3$, $\mathcal{A}(x,y) = (x, y-x, -y)$, $e_1 = (1,-3)$,
   $e_2 = (1,2)$, $h_1 = (0,-1,0)$, $h_2 = (-1,1,0)$, $h_3 = (2,0,1)$;
   2) $V = \mathbb{R}^3$, $W = \mathbb{R}^2$, $\mathcal{A}(x,y,z) = (x-y, x+z)$, $e_1 = (1,0,-2)$,
   $e_2 = (0,0,-1)$, $e_3 = (-1,2,1)$, $h_1 = (3,2)$, $h_2 = (-1,5)$;
   3) $V = \mathbb{R}^3$, $W = \mathbb{R}^2$, $\mathcal{A}(x,y,z) = (x+y, y-z)$, $e_1 = (-1,2,0)$,
   $e_2 = (0,-1,0)$, $e_3 = (2,-1,1)$, $h_1 = (3,-2)$, $h_2 = (5,1)$.

5*. Рассмотрим комплексное векторное пространство $V_n \in \mathbb{C}$.

   Определение 4.8. Числовая функция $f : V_n \to \mathbb{C}$ называется *полулинейной функцией* на пространстве V_n, если:

   $$f(u_1 + u_2) = f(u_1) + f(u_2) \quad \forall\, u_1, u_2 \in V_n;$$
   $$f(\alpha \cdot u) = \overline{\alpha} f(u) \quad \forall\, u \in V_n, \quad \forall\, \alpha \in .$$

Показать, что полулинейные функции на $V_n \in \mathbb{C}$ образуют пространство V_n^*, называемое *сопряженным пространством*. Определение сопряженного отображения $\mathcal{A}^* : V_m^* \mapsto V_n^*$ комплексных векторных пространств не отличается от определения 4.6 сопряженного отображения действительных пространств (с. 117). Все рассмотренные свойства сопряженного отображения сохраняются с небольшими изменениями, вызванными заменой линейных функций на полулинейные.

Доказать следующую теорему.

Теорема 4.8. *Пусть $\mathcal{A} : V_n \to V_m$ — линейное отображение, E и H — базисы пространств V_n и V_m, соответственно, $E^\perp$ и $H^\perp$ — биортогональные базисы пространств V_n^* и V_m^*. Тогда если отображение $\mathcal{A}$ в базисах E и H имеет матрицу $A_E^H = A$, то сопряженное отображение $\mathcal{A}^*$ в биортогональных базисах V_n^* и V_m^* имеет матрицу $A_{E^\perp}^{H^\perp} = A^*$.*

6*. Рассмотрим частный случай отображения — преобразование.

Определение 4.9. Пусть $\mathcal{A} : V_n \mapsto V_n$ — линейное преобразование пространства V_n. Тогда $\mathcal{A}^* : V_n^* \mapsto V_n^*$ — преобразование пространства V_n^*, определяемое формулой

$$\langle \mathcal{A}^*(g) \,|\, u\rangle = \langle g \,|\, \mathcal{A}(u)\rangle, \forall f \in V_n^*, \tag{4.39}$$

называется *сопряженным* преобразованию $\mathcal{A}$.

1) Пусть преобразование $\mathcal{A}$ в базисе E пространства V_n имеет матрицу A. Определить, какой вид имеет матрица сопряженного преобразования $\mathcal{A}^*$.

2) Пусть λ_i — собственные значения преобразования $\mathcal{A}$ кратности k_i. Определить собственные значения и их кратности для сопряженного преобразования $\mathcal{A}^*$.

3) Пусть базис E пространства V_n состоит из собственных векторов преобразования $\mathcal{A} : V_n \mapsto V_n$. Доказать, что его биортогональный базис $E^\perp$ состоит из собственных векторов преобразования $\mathcal{A}^*$.

4) Пусть λ — собственное значение преобразования $\mathcal{A}$, соответствующее собственному вектору $u \in V_n$, μ — собственное значение сопряженного преобразования $\mathcal{A}^*$, соответствующее собственному вектору $f \in V_n^*$. Доказать, что если $\lambda \neq \mu$, то $\langle f \,|u\,\rangle = 0$.

Вопросы для повторения

1. Дайте определение линейной функции.
2. Дайте определение сопряженного пространства.
3. Дайте определение скалярного произведения вектора и ковектора.
4. Элементами какого пространства являются ковекторы?
5. Перечислите свойства скалярного произведения.
6. Какие вектор и ковектор называются ортогональными?
7. Что такое координатные функционалы?
8. Дайте определение сопряженного базиса.
9. Приведите формулу, выражающую скалярное произведение вектора и ковектора через их координаты в произвольном базисе. Как изменится эта формула, если рассматривать биортогональные базисы?
10. В каком пространстве лежит ортогональное дополнение подпространства $V_k \subset V_n$?
11. Чему равна размерность ортогонального дополнения подпространства V_k из пространства V_n?

Необходимо усвоить

Основные определения

- линейная функция
- сопряженное пространство
- ковектор

- сопряженное отображение
- биортогональные базисы
- сопряженный базис
- ортогональное дополнение

Формулировку и трактовку теорем

- о базисных функционалах
- об ортогональном дополнении
- о сопряженном отображении
- о матрице сопряженного отображения

Основные умения и навыки

- строить биортогональный базис для заданного базиса
- находить матрицу отображения, сопряженного к заданному отображению
- находить ортогональное дополнение к заданному подпространству
- находить отображение, сопряженное к заданному линейному отображению

Глава 5

Унитарные и нормальные матрицы

Изучение свойств линейных отображений естественным образом приводит к построению и изучению свойств матриц специального вида. Ниже рассматриваются специальные классы числовых матриц, широко используемых как в теоретических, так и прикладных исследованиях.

Рекомендуемая литература:
[3]*, [5], [9], [10], [13], [15]

Цель изучения

Освоить понятие и основные свойства класса унитарных и нормальных матриц, а также матричных преобразований с их применением, получить навыки в применении нормальных и унитарных преобразований для алгебраических вычислений.

Следует повторить

- определение нормированного вектора

- определение линейно независимых и линейно зависимых векторов
- определение подобных матриц
- инварианты преобразования подобия
- определение собственного вектора и собственного значения
- определение спектра
- понятие базиса векторного пространства, его свойства
- скалярное произведение и его свойства
- диагонализуемая матрица
- треугольная матрица, её свойства

5.1. Определение и свойства унитарных матриц

Система векторов $x_1, ., x_k \in \mathbb{C}^n$ называется *ортогональной,* если

$$\langle x_i, x_j \rangle = x_i^* x_j = 0,\ \forall i, j,\ 1 \le i < j \le k.$$

Если к тому же сами векторы нормированы, т.е.

$$\langle x_i, x_i \rangle = x_i^* x_i = 1,\ i = \overline{1, k},$$

то такая система называется *ортонормированной.*

Теорема 5.1. *Любая ортонормированная система линейно независима.*

Доказательство. Рассмотрим для ортонормированной системы векторов $\{x_1, \ldots, x_n\}$ равенство

$$\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_n x_n = 0$$

и покажем, что оно возможно лишь при $\alpha_1 = \alpha_2 = \ldots = \alpha_n = 0$. Умножая равенство справа на сопряженное

$$\bar{\alpha}_1 x_1^* + \ldots + \bar{\alpha}_n x_n^* = 0^*,$$

получим

$$0 = \sum_{i,j=1}^{k} \alpha_i \alpha_j x_i^* x_j \overset{x_i^* x_j = 0,\, i \ne j}{=\!=\!=\!=\!=\!=} \sum_{i=1}^{k} |\alpha_i|^2 x_i^* x_i \overset{x_i^* x_i = 1}{=\!=\!=\!=} \sum_{i=1}^{k} |\alpha_i|^2,$$

что возможно только при $\alpha_i = 0\ \forall i = \overline{1, n}$, откуда и вытекает линейная независимость системы $\{x_1, \ldots, x_n\}$. □

Произвольную линейно независимую систему векторов можно преобразовать к ортонормированной, имеющей ту же линейную оболочку, что и исходная система. Такое преобразование можно выполнить, используя процесс *ортогонализации Грама—Шмидта* (см.[7, стр. 300]; [10, стр. 28]).

Пусть $\{x_1, \ldots, x_n\}$ — система из n линейно независимых векторов в комплексном векторном пространстве и $\{z_1, \ldots, z_n\}$ — искомая ортонормированная система. Векторы z_i вычисляются рекуррентно по следующим формулам:

$$\begin{aligned} y_1 &= x_1, & z_1 &= \tfrac{y_1}{\|y_1\|}, \\ y_k &= x_k - \sum_{j=1}^{k-1} \langle x_k, z_j \rangle z_j, & z_k &= \tfrac{y_k}{\|y_k\|},\ k = \overline{2, n}, \end{aligned} \tag{5.1}$$

где $\|x\|$ — евклидова длина вектора x.

Упражнение 5.1. Доказать, что система векторов (5.1) ортонормирована.

Пример 5.1 (Ортогонализация Грама—Шмидта). Ортонормируем систему линейно независимых векторов

$$\{x_1^T,\ x_2^T,\ x_3^T\} = \{[-1, 2, 3, 0],\ [0, 1, 2, 1],\ [2, -1, -1, 1]\}.$$

$$y_1 = x_1 = \begin{bmatrix} -1 \\ 2 \\ 3 \\ 0 \end{bmatrix}, \quad \|y_1\| = \sqrt{14}, \quad z_1 = \frac{y_1}{\sqrt{14}} = \frac{1}{\sqrt{14}} \begin{bmatrix} -1 \\ 2 \\ 3 \\ 0 \end{bmatrix};$$

$$y_2 = x_2 - \langle x_2, z_1 \rangle z_1 = \begin{bmatrix} 0 \\ 1 \\ 2 \\ 1 \end{bmatrix} - \left\langle \begin{bmatrix} 0 \\ 1 \\ 2 \\ 1 \end{bmatrix}, \frac{1}{\sqrt{14}} \begin{bmatrix} -1 \\ 2 \\ 3 \\ 0 \end{bmatrix} \right\rangle \times$$

$$\times \frac{1}{\sqrt{14}} \begin{bmatrix} -1 \\ 2 \\ 3 \\ 0 \end{bmatrix} = \frac{1}{7} \begin{bmatrix} 4 \\ -1 \\ 2 \\ 7 \end{bmatrix}, \quad \|y_2\| = \sqrt{\frac{10}{7}},$$

$$z_2 = \sqrt{\frac{7}{10}}\, y_2 = \frac{1}{\sqrt{70}} \begin{bmatrix} 4 \\ -1 \\ 2 \\ 7 \end{bmatrix};$$

$$y_3 = x_3 - \langle x_3, z_2 \rangle z_2 - \langle x_3, z_1 \rangle z_1 =$$

$$= \begin{bmatrix} 2 \\ -1 \\ -1 \\ 1 \end{bmatrix} - \left\langle \begin{bmatrix} 2 \\ -1 \\ -1 \\ 1 \end{bmatrix}, \frac{1}{\sqrt{70}} \begin{bmatrix} 4 \\ -1 \\ 2 \\ 7 \end{bmatrix} \right\rangle \times \frac{1}{\sqrt{70}} \begin{bmatrix} 4 \\ -1 \\ 2 \\ 7 \end{bmatrix} -$$

$$- \left\langle \begin{bmatrix} 2 \\ -1 \\ -1 \\ 1 \end{bmatrix}, \frac{1}{\sqrt{14}} \begin{bmatrix} -1 \\ 2 \\ 3 \\ 0 \end{bmatrix} \right\rangle \times \frac{1}{\sqrt{14}} \begin{bmatrix} -1 \\ 2 \\ 3 \\ 0 \end{bmatrix} = \frac{1}{10} \begin{bmatrix} 7 \\ 2 \\ 1 \\ -4 \end{bmatrix},$$

$$\|y_3\| = \sqrt{\frac{7}{10}}, \quad z_3 = \sqrt{\frac{10}{7}} \times y_3 = \frac{1}{\sqrt{70}} \begin{bmatrix} 7 \\ 2 \\ 1 \\ -4 \end{bmatrix}.$$

Заметим, что на любом шаге k процесса Грама—Шмидта ортонормированные векторы $z_1, \ldots, z_k$ выражаются в виде линейной комбинации только первых k исходных линейно независимых векторов $x_1, \ldots, x_k$, т.е. существуют такие числа t_{ij}, $i = \overline{1, j}$, что

$$z_j = t_{1j}x_1 + t_{2j}x_2 + \cdots + t_{jj}x_j,\ j = \overline{1, k}. \tag{5.2}$$

Справедливо и обратное.

Упражнение 5.2. Найти представление коэффициентов r_{ij} линейной комбинации, выражающей векторы x_j через z_i.

В примере 5.1, например, имеем:

$$\begin{aligned} x_1 &= \sqrt{14}\, z_1, \\ x_2 &= \frac{8}{\sqrt{14}}\, z_1 + \sqrt{\frac{10}{7}} z_2, \\ x_3 &= -\frac{7}{\sqrt{14}}\, z_1 + \frac{14}{\sqrt{70}} z_2 + \frac{\sqrt{70}}{10} z_3. \end{aligned}$$

Процесс Грама—Шмидта можно применять к любой конечной или счетной (не обязательно линейно независимой) системе векторов (см. ниже доказательство теоремы о QR-разложении).

Определение 5.1. Матрица $U \in \mathbb{C}^{n \times n}$ называется *унитарной,* если $U^*U = E$. Если к тому же $U \in \mathbb{R}^{n \times n}$, то U называется *ортогональной.*

Пример 5.2 (Унитарная матрица). Пусть $U(\theta; i, j)$ имеет вид

$$\begin{array}{cl}
\left[\begin{array}{c|ccccc|c}
\begin{matrix}1&&\\&\ddots&\\&&1\end{matrix} & & & 0 & & & \\ \hline
 & \cos\theta & 0 & \dots & 0 & \sin\theta & \\
 & 0 & 1 & & 0 & 0 & \\
0 & \vdots & & \ddots & & \vdots & 0 \\
 & 0 & 0 & & 1 & 0 & \\
 & -\sin\theta & 0 & \dots & 0 & \cos\theta & \\ \hline
 & & & 0 & & & \begin{matrix}1&&\\&\ddots&\\&&1\end{matrix}
\end{array}\right] & \begin{array}{l} i\text{-я строка} \\ \\ \\ j\text{-я строка} \end{array} \\
\begin{array}{ccc} \qquad & i\text{-й столбец} \qquad\qquad & j\text{-й столбец} \end{array} &
\end{array},$$

где θ — вещественный параметр. Эта матрица отличается от единичной лишь элементами в позициях (i, i) и (j, j), которые заменяются на $\cos\theta$, и в позициях (i, j) и (j, i), которые заменяются соответственно на $\sin\theta$ и $-\sin\theta$. Матрица $U(\theta; i, j)$ является унитарной (ортогональной) матрицей из $\mathbb{R}^{n\times n}$ для любой пары индексов $1 \le i < j \le n$ и любой величины угла θ. Так, например, при $n = 2$ имеем:

$$U^*U = U^TU = \begin{bmatrix} \cos\,\Theta & -\sin\,\Theta \\ \sin\,\Theta & \cos\,\Theta \end{bmatrix} \begin{bmatrix} \cos\,\Theta & \sin\,\Theta \\ -\sin\,\Theta & \cos\,\Theta \end{bmatrix} =$$
$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Теорема 5.2 (о критериях унитарности). *Следующие предложения эквивалентны для матрицы $U \in \mathbb{C}^{n\times n}$:*

a) U унитарная;

b) U невырожденная и $U^* = U^{-1}$;

c) $UU^* = E_n$;

d) U^* унитарная;

e) столбцы в U образуют ортонормированную систему;

f) строки в U образуют ортонормированную систему;

g) для любого вектора $x \in \mathbb{C}^n$ справедливо $x^*x = (Ux)^*(Ux)$ (т.е. унитарные матрицы — изометричные).

Доказательство. a) влечет b), т.к. матрица U^{-1} (если она существует) — это единственная матрица, при умножении на которую слева получается E. Из определения унитарности следует, что U^* именно такая матрица: $UU^* = UU^{-1} = E$. Значит, из b) следует c).

Так как c) — это определение унитарности для U^*, то из c) следует d).

Обращение импликаций проводится аналогично. Таким образом $a), b), c)$ и d) эквивалентны.

Обозначим u_i — i-й столбец матрицы U. Из правил матричного умножения следует, что если $U^*U = E$, то $u_i{}^*u_j = \delta_{ij}$, где

$$\delta_{ij} = \begin{cases} 0, & i \ne j, \\ 1, & i = j, \end{cases}$$

— символ Кронекера. Последнее равносильно тому, что u_i, u_j ортонормированные. Следовательно, a) равносильно e). Аналогично доказывается равносильность d) и f).

Докажем, что из a) следует g). Возьмем $y = Ux$. Тогда

$$y^*y = x^*U^*Ux = x^*x.$$

Докажем, что из g) следует a). Рассмотрим сначала случай $n = 2$. Рассмотрим произвольную матрицу U, для которой выполняется g) для $\forall x \in \mathbb{C}^n$. Возьмем

$$x = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Имеем:

$$1 = x^*x = y^*y = x^*U^*Ux$$

— это элемент матрицы U^*U в позиции $[1,1]$. Аналогично полагая

$$x = \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

получим $U^*U_{(22)} = 1$. Т.е. U^*U имеет вид

$$\begin{bmatrix} 1 & a \\ \bar{a} & 1 \end{bmatrix},$$

где a — скалярное произведение 1-го и 2-го столбца, $\bar{a}$ — 2-го и 1-го столбца. Положим

$$x = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Тогда из g) следует

$$2 = x^*x = y^*y = x^*U^*Ux = \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & a \\ \bar{a} & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} =$$

$$= \begin{bmatrix} 1+\bar{a} & a+1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 2 + (\bar{a} + a) \Rightarrow a + \bar{a} = 2\operatorname{Re} a = 0.$$

Рассматривая

$$x = \begin{bmatrix} 1 \\ i \end{bmatrix},$$

будем иметь:

$$2 = x^*x = y^*y = x^*U^*Ux = 2 + i(a - \bar{a}) \Rightarrow a - \bar{a} = 2i\operatorname{Im} a = 0.$$

Таким образом $\bar{a} + a = 2\operatorname{Re} a = 0$, $a - \bar{a} = 2i\operatorname{Im} a = 0 \Rightarrow a = 0$. Следовательно, при $n = 2$ в случае выполнения g) имеем $UU^* = E$, т.е. верно a).

Рассмотрим теперь случай $n > 2$. Пусть $U \in \mathbb{C}^{n\times n}$ — матрица, для которой выполняется g). Положим $A = U^*U$. Возьмем $x \in \mathbb{C}^n$, в котором все компоненты нулевые, кроме i-й и j-й ($i \le j$). Тогда

$$yy^* = x^*Ax = [\bar{x}_i, \bar{x}_j]\ A_{\{i,j\}} \begin{bmatrix} x_i \\ x_j \end{bmatrix},$$

где $A_{\{i,j\}}$ — главная подматрица матрицы A в строках i, j и столбцах i, j. Из доказанного выше при $n = 2$ следует, что $A_{\{i,j\}} = E_2 \in \mathbb{C}^{2\times 2}$. Так как i, j — произвольные, то в A любая главная подматрица порядка 2 совпадает с E_2. Единственная матрица, обладающая таким свойством — это матрица E_n.

Случай $n = 1$ очевиден. □

Теорема 5.3 (о QR-разложении). *Если $A \in \mathbb{C}^{n\times m}$, $n \ge m$, то существуют матрица $Q \in \mathbb{C}^{n\times m}$ с ортонормированными столбцами и верхняя треугольная матрица $R \in \mathbb{C}^{m\times m}$, такие, что*

$$A = QR.$$

Если $m = n$, то Q унитарна.

Доказательство. Если $A \in \mathbb{C}^{n\times m}$ и $\operatorname{rank} A = m$, то QR−разложение матрицы A есть не что иное, как матричная запись результата применения процесса Грама—Шмидта к столбцам матрицы A, образующим в $\mathbb{C}^n$ линейно независимую систему.

* Пусть столбцы матрицы $A = [a_1, \ldots, a_m]$ линейно зависимы. Обобщим алгоритм Грама—Шмидта на этот случай.

Для тех k, $k = \overline{1, m}$, для которых в процессе ортогонализации Грама—Шмидта получается $y_k = 0$ (т.е. a_k есть линейная комбинация $a_1, \ldots, a_{k-1}$), полагаем $z_k = 0$. В противном случае $z_k = \frac{y_k}{\|y_k\|}$ (как в обычном процессе Грама—Шмидта).

Векторы $z_1, \ldots, z_m$ составляют ортогональную систему, каждый элемент которого нормированный или нулевой.

Каждый вектор z_j — это линейная комбинация векторов $a_1, \ldots, a_j$, и обратно. Следовательно, найдутся числа r_{ij} (см. упражнение 5.2) такие, что

$$a_j = \sum_{i=1}^{j} r_{ij} z_i, \; j = \overline{1, m}. \tag{5.3}$$

Положим

$$\begin{aligned} &r_{ij} = 0, \; i > j, \\ &r_{ij} = 0, \; j = \overline{1, m}, \text{ если } \quad z_i = 0. \end{aligned} \tag{5.4}$$

Таким образом, исходя из $a_1, \ldots, a_m$, с помощью описанной процедуры определим верхнюю треугольную матрицу

$$R = [r_{ij}] \in \mathbb{C}^{n \times n}$$

и векторы $z_1, \ldots, z_m$. Матрица

$$Q \equiv [z_1, \ldots, z_m] \in \mathbb{C}^{n \times m}$$

имеет ортогональные столбцы (некоторые могут быть нулевые), и в силу (5.3) $A = QR$.

Если $m = n$ и $\operatorname{rank} A = m$ (т.е. A невырожденная), то Q — унитарная (по свойству e) теоремы 2) и все диагональные элементы матрицы $R = Q^* A$ отличны от нуля. В этом случае так как матрица R — верхняя треугольная, вектор z_1 есть кратное вектора a_1 и при $i = 2, 3, \ldots, m$ вектор z_i лежит в одномерном пространстве, которое является ортогональным дополнением линейной оболочки векторов $a_1, \ldots, a_{i-1}$ в линейной оболочке векторов $a_1, \ldots, a_i$. Следовательно, каждый вектор z_i определяется однозначно с точностью до скалярного сомножителя, по модулю равного 1. Поэтому, заменяя R на R':

$$R' = \operatorname{diag}\left[\frac{|r_{11}|}{r_{11}}, \ldots, \frac{|r_{mm}|}{r_{mm}}\right] R$$

и Q на Q':

$$Q' = Q \times \operatorname{diag}\left[\frac{r_{11}}{|r_{11}|}, \ldots, \frac{r_{mm}}{|r_{mm}|}\right],$$

получаем то единственное разложение, о котором идет речь в утверждении теоремы.

Если столбцы в A зависимы, то возьмем (ортонормированное) множество ненулевых столбцов в Q и дополним его до ортонормированного базиса в $\mathbb{C}^n$. Новые векторы, полученные таким способом, обозначим $z_1, \ldots, z_p$. Теперь заменим первый нулевой столбец в Q на z_1, второй на z_2 и т.д. Полученную матрицу обозначим через Q'. Она имеет ортонормированные столбцы, и $QR = Q'R$, потому что новые столбцы в Q' соответствуют нулевым строкам в R (см. (5.4)). Таким образом, $A = Q'R$ — разложение нужного вида. □

Пример 5.3 (QR-разложение). Построим QR-разложение для матрицы

$$A = \begin{bmatrix} -1 & 0 & 2 \\ 2 & 1 & -1 \\ 3 & 2 & -1 \\ 0 & 1 & 1 \end{bmatrix}.$$

Используем для этого процесс ортогонализации Грама—Шмидта. Обозначим столбцы матрицы A

$$x_1^T = (-1, 2, 3, 0), \quad x_2^T = (0, 1, 2, 1), \quad x_3^T = (2, -1, -1, 1).$$

Как показано в примере 5.1, в результате процесса ортогонализации Грама—Шмидта система векторов $\{x_1, \; x_2, \; x_3\}$ преобразуется в ортонормированную систему векторов $\{z_1, \; z_2, \; z_3\}$, где

$$z_1 = \frac{1}{\sqrt{14}} \begin{bmatrix} -1 \\ 2 \\ 3 \\ 0 \end{bmatrix}, \; z_2 = \frac{1}{\sqrt{70}} \begin{bmatrix} -4 \\ -1 \\ 2 \\ 7 \end{bmatrix}, \; z_3 = \frac{1}{\sqrt{70}} \begin{bmatrix} -7 \\ 2 \\ 1 \\ -4 \end{bmatrix},$$

причем

$$\begin{aligned} x_1 &= \sqrt{14}\, z_1, \\ x_2 &= \frac{8}{\sqrt{14}} z_1 + \frac{10}{\sqrt{70}} z_2, \\ x_3 &= -\frac{7}{\sqrt{14}} z_1 + \frac{14}{\sqrt{70}} z_2 + \frac{\sqrt{70}}{10} z_3. \end{aligned}$$

Образуем из векторов $z_1,\ z_2,\ z_3$ матрицу

$$Q = [z_1,\ z_2,\ z_3] = \begin{bmatrix} -\frac{1}{\sqrt{14}} & \frac{4}{\sqrt{70}} & \frac{7}{\sqrt{70}} \\ \frac{2}{\sqrt{14}} & -\frac{1}{\sqrt{70}} & \frac{2}{\sqrt{70}} \\ \frac{3}{\sqrt{14}} & \frac{2}{\sqrt{70}} & \frac{1}{\sqrt{70}} \\ 0 & \frac{7}{\sqrt{70}} & -\frac{4}{\sqrt{70}} \end{bmatrix}.$$

Тогда из равенств, связывающих x_i и z_i, $i = 1, 2, 3$, имеем:

$$A = [x_1,\ x_2,\ x_3] = QR, \text{где} \quad R = \begin{bmatrix} \sqrt{14} & \frac{8}{\sqrt{14}} & -\frac{7}{\sqrt{14}} \\ 0 & \frac{10}{\sqrt{70}} & \frac{14}{\sqrt{70}} \\ 0 & 0 & \frac{\sqrt{70}}{10} \end{bmatrix}.$$

5.2. Унитарное подобие

Так как для унитарной матрицы $U^* = U^{-1}$, то преобразование $A \to U^*AU$, определенное на $\mathbb{C}^{n\times n}$, является подобием, которое называется *унитарным подобием.*

Определение 5.2. Матрица $B \in \mathbb{C}^{n\times n}$ называется *унитарно подобной* матрице $A \in \mathbb{C}^{n\times n}$, если найдется унитарная матрица $U \in \mathbb{C}^{n\times n}$, такая, что $B = U^*AU$. Если U можно выбрать вещественной (а значит, ортогональной), то B называется *ортогонально подобной* матрице A.

Рассмотрим два специальных типа унитарных матриц, которые осуществляют преобразования унитарного подобия, весьма важные для вычисления собственных значений.

Пример 5.4 (Плоские вращения). Матрица $U(\theta; i, j)$ из примера 5.2 осуществляет вращение (на угол θ) в плоскости координат i, j. Заметим, что если матрица умножается слева на $U(\theta; i, j)$, то в ней изменяются только i-я и j-я строки, а если она умножается справа, то изменяются только i-й и j-й столбцы. Таким образом, при переходе к унитарно подобной матрице, осуществляемом с помощью $U(\theta; i, j)$, происходит изменение только строк и столбцов с номерами i и j. Унитарное подобие, осуществляемое посредством плоских вращений, используется при вычислении собственных значений.

Пример 5.5* (Преобразования Хаусхолдера). Возьмем произвольный ненулевой вектор $w \in \mathbb{C}^n$ и образуем матрицу

$$U_w = E - tww^* \ (U_w \in \mathbb{C}^{n\times n}),$$

где $t = 2(w^*w)^{-1}$. Заметим, что w^*w — это положительный скаляр, $ww^* \in \mathbb{C}^{n\times n}$ — матрица. Если вектор w был нормирован ($w^*w = 1$), то t должно быть равно 2, а матрица U_w должна иметь вид $U_w = E - 2ww^*$. Часто образуют матрицу U_w, выбирая заранее именно нормированный вектор w.

Любая матрица U_w называется *преобразованием Хаусхолдера.*

Теорема 5.4. *Для унитарно подобных матриц A и B из $\mathbb{C}^{n\times n}$ имеет место равенство*

$$\sum_{i,j=1}^{n} |a_{ij}|^2 = \sum_{i,j=1}^{n} |b_{ij}|^2.$$

Доказательство вытекает из следующей цепочки равенств, получаемых с учетом леммы 2.5 и инвариантности следа матрицы относительно преобразования подобия:

$$\begin{aligned} &\sum_{i,j=1}^{n} |b_{ij}|^2 = \operatorname{tr} B^*B = \operatorname{tr} (U^*AU)^* U^*AU = \\ &= \operatorname{tr} U^*A^*UU^*AU = \operatorname{tr} U^*A^*AU = \operatorname{tr} A^*A = \sum_{i,j=1}^{n} |a_{ij}|^2. \end{aligned}$$

$\square$

Следующая теорема показывает, что произвольная матрица $A \in \mathbb{C}^{n\times n}$ унитарно подобна какой-то верхней треугольной матрице (а также нижней треугольной матрице).

Теорема 5.5 (Шура об унитарной триангуляризации). *Пусть задана матрица $A \in \mathbb{C}^{n\times n}$ и зафиксирован какой-то порядок ее собственных значений $\lambda_1, \ldots, \lambda_n$. Тогда существует унитарная матрица $U \in \mathbb{C}^{n\times n}$, такая, что*

$$U^*AU = T = [t_{ij}] -$$

верхняя треугольная матрица с диагональными элементами $t_{ii} = \lambda_i$, $i = \overline{1,n}$. Кроме того, если $A \in \mathbb{R}^{n\times n}$ и все ее собственные значения вещественны, то U можно выбрать ортогональной.

Другими словами, любая комплексная матрица унитарно подобна треугольной матрице.

Доказательство носит алгоритмический характер и сводится к последовательности однотипных редукций.

Пусть $x^{(1)}$ — нормированный собственный вектор матрицы A, отвечающий собственному значению λ_1: $Ax^{(1)} = \lambda_1 x^{(1)}$. Вектор $x^{(1)}$ ненулевой и в $\mathbb{C}^n$ его можно дополнить до базиса

$$x^{(1)}, y^{(2)}, \ldots, y^{(n)}.$$

Применим процесс ортогонализации Грама—Шмидта и перейдем в $\mathbb{C}^n$ к ортонормированному базису

$$x^{(1)}, z^{(2)}, \ldots, z^{(n)}.$$

Из этих векторов (слева-направо) образуем унитарную матрицу U_1:

$$U_1 = \left[x^{(1)}, z^{(2)}, \ldots, z^{(n)}\right].$$

Рассмотрим матрицу AU_1:

$$AU_1 = A\left[x^{(1)}, z^{(2)}, \ldots, z^{(n)}\right] = \left[Ax^{(1)}, Az^{(2)}, \ldots, Az^{(n)}\right] =$$
$$= \left[\lambda_1 x^{(1)}, Az^{(2)}, \ldots, Az^{(n)}\right]$$

и матрицу

$$\tilde{A}_1 = U_1^*(AU_1) = \begin{bmatrix} x^{(1)*} \\ z^{(2)*} \\ \ldots \\ z^{(n)*} \end{bmatrix} \begin{bmatrix} \lambda_1 x^{(1)} & Az^{(2)} & \ldots & Az^{(n)} \end{bmatrix} =$$
$$= \begin{bmatrix} x^{(1)*}\lambda x^{(1)} & * \\ 0 & A_1 \end{bmatrix} = \begin{bmatrix} \lambda_1 & * \\ 0 & A_1 \end{bmatrix} = \tilde{A}_1.$$

Так как $\tilde{A}_1 = U_1^*AU_1$ подобна A, а значит, имеет те же собственные значения, то собственные значения матрицы $A_1 \in \mathbb{C}^{(n-1)\times(n-1)}$ равны $\lambda_2, \ldots, \lambda_n$.

Пусть теперь $x^{(2)} \in \mathbb{C}^{(n-1)}$ есть собственный вектор для A_1, отвечающий λ_2. Далее все повторяется. То есть для некоторой унитарной матрицы

$$U_2 \in \mathbb{C}^{(n-1)\times(n-1)}: \quad U_2 = \left[x^{(2)}, w^{(3)}, \ldots, w^{(n)}\right]$$

получаем

$$U_2^*A_1U_2 = \begin{bmatrix} \lambda_2 & * \\ 0 & A_2 \end{bmatrix}.$$

Сформируем матрицу

$$V_2 = \begin{bmatrix} 1 & 0 \\ 0 & U_2 \end{bmatrix}.$$

Легко проверить, что матрица V_2 унитарна. Далее, произведение двух унитарных матриц U_1V_2 — унитарная матрица, поскольку

$$(U_1V_2)^*(U_1V_2) = V_2^*U_1^*U_1V_2 = V_2^*V_2 = E.$$

При этом

$$V_2^* U_1^* A U_1 V_2 = \begin{bmatrix} \lambda_1 & * & * \\ 0 & \lambda_2 & * \\ 0 & 0 & A_2 \end{bmatrix}.$$

Продолжаем редукцию и находим унитарные матрицы $U_i \in \mathbb{C}^{(n-i+1)\times(n-i+1)}$, $i = \overline{1, n-1}$ и унитарные матрицы $V_i \in \mathbb{C}^{n\times n}$. Матрица $U = U_1 V_2 V_3 \dots V_{n-1}$ унитарная, и она обеспечивает треугольный вид матрицы $U^* A U$.

Если все собственные значения матрицы $A \in \mathbb{R}^{n\times n}$ вещественны, то отвечающие им собственные векторы тоже можно выбрать вещественными, т.е. все шаги реализуются в вещественной арифметике, что и доказывает заключительное утверждение. □

Замечание 5.1. В формулировке теоремы Шура можно говорить не о верхних треугольных, а о нижних треугольных матрицах. При этом унитарная эквивалентность будет осуществляться другой матрицей U.

Замечание 5.2. В теореме Шура неоднозначно определяются и U, и T. Это связано с выбором векторов $z^{(i)}, w^{(i)}$ и т.п., порядком расстановки собственных чисел. Кроме того, то, что находится выше главной диагонали для унитарно подобных верхних треугольных матриц может сильно различаться. Например, следующие матрицы T_1 и T_2 — унитарно подобны с унитарной матрицей преобразования U:

$$T_1 = \begin{bmatrix} 1 & 1 & 4 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{bmatrix}, \ T_2 = \begin{bmatrix} 2 & -1 & 3\sqrt{2} \\ 0 & 1 & \sqrt{2} \\ 0 & 0 & 3 \end{bmatrix}, \ U = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

Применение теоремы Шура позволяет установить «почти диагонализуемость» произвольной матрицы.

Следствие 5.1*. *Пусть $A = [a_{ij}] \in \mathbb{C}^{n\times n}$. Для $\forall\, \varepsilon > 0$ существует матрица $A(\varepsilon) = [a_{ij}(\varepsilon)] \in \mathbb{C}^{n\times n}$, имеющая n различных собственных значений (и, значит, диагонализуемая) и такая, что*

$$\sum_{i,j=1}^{n} |a_{ij} - a_{ij}(\varepsilon)|^2 < \varepsilon.$$

Другими словами, для любой матрицы существует сколь угодно близкая к ней диагонализуемая матрица.

Доказательство. Пусть матрица $U \in \mathbb{C}^{n\times n}$ унитарная и такая, что $U^* A U = T$ — верхняя треугольная матрица. Образуем матрицу $\mathrm{E} = \mathrm{diag}\,[e_1, \dots, e_n]$, при этом числа $e_1, \dots, e_n$ выберем таким образом, чтобы для всех i выполнялось неравенство $|e_i| < \left[\frac{\varepsilon}{n}\right]^{\frac{1}{2}}$ и, кроме того, числа $t_{11} + e_1, \dots, t_{nn} + e_n$ были различны. Матрица $T + \mathrm{E}$ имеет n различных собственных значений $t_{11} + e_1, \dots, t_{nn} + e_n$.

Рассмотрим матрицу $A(\varepsilon) = A + U\mathrm{E}U^*$. Покажем, что она унитарно подобна $T + \mathrm{E}$:

$$U^* A(\varepsilon) U = U^* (A + U\mathrm{E}U^*) U = U^* A U + U^* U \mathrm{E} U^* U = T + \mathrm{E},$$

поэтому она тоже имеет n различных собственных значений.

Рассмотрим две унитарно подобные матрицы: E и $A(\varepsilon) - A = U\mathrm{E}U^*$. По теореме 5.4 имеем:

$$\sum_{i,j=1}^{n} |a_{ij} - a_{ij}(\varepsilon)|^2 = \sum_{i,j=1}^{n} |e_i|^2 < n \left[\frac{\varepsilon}{n}\right] = \varepsilon.$$ □

Упражнение 5.3. Показать, что действительно возможно выбрать требуемые при доказательстве теоремы числа $e_1, \dots, e_n$.

Следствие 5.2*. *Пусть $A \in \mathbb{C}^{n\times n}$. Для $\forall\, \varepsilon > 0$ существует невырожденная матрица $S_\varepsilon \in \mathbb{C}^{n\times n}$, такая, что матрица*

$$S_\varepsilon^{-1} A S_\varepsilon = T_\varepsilon = [t_{ij}(\varepsilon)]$$

— верхняя треугольная и $|t_{ij}(\varepsilon)| < \varepsilon$ для $1 \le i < j \le n$.

Другими словами, любая матрица подобна верхней треугольной матрице с произвольно малыми внедиагональными элементами.

Доказательство. Во-первых, по теореме Шура существуют унитарная матрица U и верхняя треугольная матрица T, такие что

$$U^*AU = T.$$

Возьмем ненулевое число α, образуем матрицу

$$D_\alpha = \operatorname{diag}\left[1, \alpha, \alpha^2, \ldots, \alpha^{n-1}\right]$$

и положим

$$t = \max_{i<j} |t_{ij}|\,.$$

Пусть $\varepsilon < 1$ (достаточно рассмотреть только этот случай). Положим

$$\begin{aligned} S_\varepsilon &= UD_\varepsilon, \quad \text{если} \quad t \le 1, \\ S_\varepsilon &= UD_{\frac{1}{t}}D_\varepsilon, \quad \text{если} \quad t > 1. \end{aligned}$$

Покажем, что в обоих случаях S_ε — искомая матрица. Действительно, если $t \le 1$, то

$$\begin{aligned} S_\varepsilon^{-1}AS_\varepsilon &= \left[UD_\varepsilon\right]^{-1}AUD_\varepsilon = D_\varepsilon^{-1}U^{-1}AUD_\varepsilon = \\ &= \operatorname{diag}\left[1, \frac{1}{\varepsilon}, \frac{1}{\varepsilon^2}, \ldots, \frac{1}{\varepsilon^{n-1}}\right] T \operatorname{diag}\left[1, \varepsilon, \varepsilon^2, \ldots, \varepsilon^{n-1}\right] = \\ &= \left[\frac{1}{\varepsilon^i}t_{ij}\varepsilon^j\right] = \left[t_{ij}\varepsilon^{j-i}\right] = \left[t_{ij}(\varepsilon)\right] = T(\varepsilon)\,, \end{aligned}$$

где

$$|t_{ij}(\varepsilon)| = \left|t_{ij}\varepsilon^{j-i}\right| = |t_{ij}|\left|\varepsilon^{j-i}\right| \le \left|\varepsilon^{j-i}\right| \overset{(i<j)}{\le} \varepsilon\,.$$

Если $t > 1$, то

$$\begin{aligned} S_\varepsilon^{-1}AS_\varepsilon &= \left[UD_{\frac{1}{t}}D_\varepsilon\right]^{-1}AUD_{\frac{1}{t}}D_\varepsilon = D_\varepsilon^{-1}D_{\frac{1}{t}}^{-1}U^{-1}AUD_{\frac{1}{t}}D_\varepsilon = \\ &= \operatorname{diag}\left[1, \frac{1}{\varepsilon}, \frac{1}{\varepsilon^2}, \ldots, \frac{1}{\varepsilon^{n-1}}\right] \operatorname{diag}\left[1, t, \ldots, t^{n-1}\right] \times \\ &\times T \operatorname{diag}\left[1, \frac{1}{t}, \ldots, \frac{1}{t^{n-1}}\right] \operatorname{diag}\left[1, \varepsilon, \varepsilon^2, \ldots, \varepsilon^{n-1}\right] = \\ &= \left[\varepsilon^{-i}t^it_{ij}t^{-j}\varepsilon^j\right] = \left[t_{ij}(\varepsilon)\right] = T(\varepsilon)\,, \end{aligned}$$

где

$$|t_{ij}(\varepsilon)| = \left|\varepsilon^{j-i}t^{i-j}t_{ij}\right| \le \left|\varepsilon^{j-i}\right| \times \left|\frac{t_{ij}}{(\max t_{ij})^{j-i}}\right| \le \varepsilon\,. \qquad \square$$

Теорема Шура позволяет весьма просто доказать следующий полезный результат.

Теорема 5.6. *Если $A \in \mathbb{C}^{n\times n}$ имеет собственные значения $\lambda_1, \ldots, \lambda_n$ (с учетом кратностей), то*

$$\det A = \prod_{i=1}^{n} \lambda_i, \quad \operatorname{tr} A = \sum_{i=1}^{n} \lambda_i.$$

Доказательство. Используя теорему Шура, запишем $U^*AU = T$. Тогда с учетом инвариантности следа и определителя матрицы относительно преобразования подобия имеем:

$$\operatorname{tr} A = \operatorname{tr} UTU^* = \operatorname{tr} T = \sum_i \lambda_i, \tag{5.5}$$

$$\det A = \det UTU^* = \det T = \prod_i \lambda_i, \tag{5.6}$$

что и завершает доказательство. $\square$

Свойства унитарных матриц

1^0. Для любых унитарных (ортогональных) матриц $U,\ V \in \mathbb{C}^{n\times n}$ произведение UV является также унитарной (ортогональной) матрицей.

2^0. Все собственные значения унитарной матрицы по модулю равны 1.

Определение 5.3. *Матрицей перестановок* называется квадратная матрица, у которой в каждой строке и каждом столбце только один элемент отличен от нуля и равен единице.

Например, при умножении заданной матрицы A на матрицу перестановок

$$P = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

слева меняются местами первая и вторая строка, а при умножении справа — первый и второй столбец.

3^0. Любая матрица перестановок — унитарная.

4^0. Определитель унитарной матрицы по модулю равен 1.

5^0. Унитарные матрицы одного порядка образуют группу по умножению.

Доказательство. 1^0. Пусть U, V — унитарные. Используя свойства операции $(^*)$, имеем:

$$(UV)^*(UV) = V^*U^*UV = E,$$

откуда следует, что UV — также унитарная.

2^0. Пусть λ — собственное значение унитарной матрицы U, x — соответствующий собственный вектор: $Ux = \lambda x$. Тогда

$$(Ux)^*(Ux) = (\lambda x)^*\lambda x \Rightarrow x^*U^*Ux = |\lambda|^2 x^*x \Rightarrow$$
$$\Rightarrow x^*x = |\lambda|^2 x^*x \Rightarrow |\lambda| = 1.$$

3^0. Проверяется непосредственно по определениям матрицы перестановок и унитарной матрицы.

4^0. Следует из 2^0 и теоремы 5.6.

5^0. Легко доказывается с использованием 1^0, ассоциативности умножения матриц, унитарности единичной матрицы и b) из теоремы 5.2. □

5.3. Нормальные матрицы

Определение 5.4. Матрица $A \in \mathbb{C}^{n\times n}$ называется *нормальной,* если она перестановочна со своей сопряженной матрицей:

$$AA^* = A^*A.$$

Пример 5.6 (Нормальные матрицы).

a) если матрица U унитарна, то $U^*U = E = UU^*$; поэтому все унитарные матрицы нормальны;

b) если $A^* = A$ (матрица A *эрмитова*), то, очевидно, что $A^*A = AA^*$. Поэтому все эрмитовы матрицы нормальные;

c) если $A = -A^*$ (матрица *косоэрмитова*), то $A^*A = -A^2 = AA^*$. Поэтому все косоэрмитовы матрицы тоже нормальны;

d) симметричные и кососимметричные матрицы являются нормальными;

e) матрица $A = \left[\begin{smallmatrix} 1 & -1 \\ 1 & 1 \end{smallmatrix}\right]$ нормальна и не относится ни к одному из перечисленных выше классов.

Свойства нормальных матриц

1^0. Блочно-треугольная нормальная матрица T является блочно-диагональной.

2^0. Диагональная матрица является нормальной.

3^0. Пусть $A \in \mathbb{C}^{n\times n}$ — произвольная матрица, $\alpha \in \mathbb{C}, \beta \in \mathbb{C}$, $|\alpha| = |\beta|$. Тогда матрица $\alpha A + \beta A^*$ является нормальной.

4^0. Пусть A — нормальная матрица. Для любого полинома $\varphi(\lambda)$ матрица $\varphi(A)$ будет нормальной.

5^0. Если A — нормальная невырожденная матрица, то A^{-1} тоже нормальная.

Доказательство. Свойство 1^0 доказывается приравниванием диагональных блоков матриц TT* и T*T.

Свойство 2^0 очевидно.

Доказательство свойства 3^0 вытекает из следующей цепочки равенств:

$$(\alpha A+\beta A^*)(\alpha A+\beta A^*)^* = (\alpha A+\beta A^*)\left(\bar\alpha A^*+\bar\beta A\right) =$$
$$= |\alpha|^2 AA^* + \alpha\bar\beta A^2 + \beta\bar\alpha (A^*)^2 + |\beta|^2 A^*A \overset{|\alpha|^2=|\beta|^2}{=\!=\!=\!=\!=}$$
$$= |\alpha|^2 A^*A + \bar\alpha\beta (A^*)^2 + \bar\beta\alpha A^2 + |\beta|^2 AA^* =$$
$$= \left(\bar\alpha A^*+\bar\beta A\right)(\alpha A+\beta A^*) = (\alpha A+\beta A^*)^*(\alpha A+\beta A^*).$$

Докажем свойство 4^0. Пусть

$$\varphi(\lambda) = a_n\lambda^n + a_{n-1}\lambda^{n-1} + \ldots + a_1\lambda + a_0.$$

Тогда

$$\varphi(A)(\varphi(A))^* = \varphi(A)\varphi(A^*) =$$
$$= (a_nA^n + \ldots + a_1A + a_0E)(a_n(A^*)^n + \ldots + a_1A^* + a_0E) =$$
$$= \sum_{i,j=0}^{n} a_ia_jA^i(A^*)^j \overset{AA^*=A^*A}{=\!=\!=\!=\!=\!=} \sum_{j,i=0}^{n} a_ja_i(A^*)^jA^i =$$
$$= \Big(a_n(A^*)^n + \ldots + a_1A^* + a_0E\Big)\Big(a_nA^n + \ldots + a_1A + a_0E\Big) =$$
$$= \varphi(A^*)\varphi(A) = (\varphi(A))^*\varphi(A).$$

Докажем 5^0. Пусть A— нормальная, т.е. $AA^* = A^*A$. Используя свойства операции обращения, имеем:

$$A^{-1}(A^{-1})^* = A^{-1}(A^*)^{-1} = (A^*A)^{-1} =$$
$$= (AA^*)^{-1} = (A^*)^{-1}A^{-1} = (A^{-1})^*A^{-1},$$

что согласно определению 5.4 доказывает, что A^{-1} — тоже нормальная. □

Определение 5.5. Матрица $A \in \mathbb{C}^{n\times n}$ называется *унитарно диагонализуемой*, если она унитарно подобна какой-либо диагональной матрице. Аналогично вводится понятие *ортогонально диагонализуемой* матрицы.

Таким образом, для унитарно диагонализуемой матрицы имеет место представление

$$A = U\Lambda U^* \tag{5.7}$$

где $\Lambda = \operatorname{diag}\{\lambda_1,\ldots,\lambda_n\}$ — диагональная, а U — унитарная матрицы, причем диагональные элементы матрицы Λ составляют спектр матрицы A.

Определение 5.6. Представление (5.7) называется *спектральным разложением* матрицы A.

Теорема 5.7 (Спектральная теорема для нормальных матриц). *Для матрицы $A = [a_{ij}] \in \mathbb{C}^{n\times n}$ с собственными значениями $\lambda_1,\ldots,\lambda_n$ следующие утверждения равносильны:*

a) A нормальна;

b) A унитарно диагонализуема;

c) $\sum\limits_{i,j=1}^{n} |a_{ij}|^2 = \sum\limits_{i=1}^{n} |\lambda_i|^2$;

d) для A существует ортонормированная система из собственных векторов.

Доказательство. a) $\Rightarrow$ b) : по теореме Шура существует унитарная матрица $V \in \mathbb{C}^{n\times n}$ и треугольная матрица $T \in \mathbb{C}^{n\times n}$, такие что

$$A = VTV^*.$$

Умножим это равенство на сопряженное. С учетом a) имеем

$$\begin{aligned} AA^* &= VTV^*VT^*V^* = VTT^*V^* = A^*A = \\ &= VT^*V^*VTV^* = VT^*TV^*, \end{aligned} \tag{5.8}$$

откуда следует

$$TT^* = T^*T,$$

т.е. T — треугольная нормальная матрица и по свойству 1^0 нормальных матриц T — диагональная.

b) $\Rightarrow$ a) : по определению унитарно диагонализуемой матрицы для матрицы A существуют унитарная матрица $U \in \mathbb{C}^{n\times n}$ и диагональная матрица $\Lambda \in \mathbb{C}^{n\times n}$, такие что

$$A = U\Lambda U^*.$$

Следовательно, с учетом свойства 2^0 нормальных матриц

$$AA^* = U\Lambda U^*U\Lambda^*U^* = U\Lambda\Lambda^*U^* = U\Lambda^*\Lambda U^* = A^*A,$$

т.е. A — нормальная.

b) $\Rightarrow$ c) следует из примения теоремы 5.4 к унитарно подобным матрицам A и Λ.

c)$\Rightarrow$ b). По теореме Шура $A = VTV^*$, где $t_{ii} = \lambda_i(A)$. Далее применяя теорему 5.4 к унитарно подобным матрицам A и T, имеем:

$$\sum_{i,j} |a_{ij}|^2 = \sum_i |t_{ii}|^2 + \sum_{i<j} |t_{ij}|^2,$$

откуда с учетом c) следует

$$\sum_{i<j} |t_{ij}|^2 = 0 \Rightarrow t_{ij} = 0 \ \ \forall i < j.$$

a) $\Rightarrow$ d) Обозначим $u_i,\ i = \overline{1,n}$ — столбцы матрицы U в разложении (5.7):

$$U = [u_1, u_2, \ldots, u_n]\,.$$

Из разложения (5.7) в силу $U^* = U^{-1}$ для унитарной матрицы U имеем $AU = U\Lambda$, что с учетом того, что матрица Λ — диагональная, равносильно $Au_i = \lambda_i u_i,\ i = \overline{1,n}$. Из последних равенств и определения собственных векторов вытекает, что $u_i,\ i = \overline{1,n}$, являются собственными векторами матрицы A. В силу утверждения e) теоремы 5.2 векторы u_i образуют ортонормированную систему, что и завершает доказательство.

d) $\Rightarrow$ a) Пусть $u_i, i = \overline{1,n}$ — система ортонормированных собственных векторов матрицы :

$$Au_i = \lambda_i u_i, i = \overline{1,n}. \tag{5.9}$$

Составим из этих векторов матрицу $U = [u_1, \ldots, u_n]$. Тогда (5.9) можно записать в матричном виде

$$AU = \Lambda U = U\Lambda,$$

откуда следует

$$U^*AU = \Lambda,$$

т.е. A — унитарно подобна диагональной, а значит, согласно b), A — нормальна. □

Следствие 5.3. *Столбцами унитарной матрицы U в спектральном разложении (5.7) нормальной матрицы являются ортонормированные собственные векторы матрицы A.*

Пример 5.7 (Спектральное разложение нормальной матрицы). Матрица A из примера 5.6 e) — нормальная, а значит, унитарно диагонализуемая. Построим спектральное разложение (5.7) матрицы A, используя следствие 5.3.

Матрица A имеет собственные значения $1-i,\ 1+i$ и соответствующие им нормированные собственные векторы

$$\begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{i}{\sqrt{2}} \end{bmatrix}, \begin{bmatrix} \frac{i}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}.$$

Диагональная матрица Λ из (5.7) с собственными значениями матрицы A на диагонали имеет вид

$$\Lambda = \begin{bmatrix} -i+1 & 0 \\ 0 & i+1 \end{bmatrix},$$

а унитарная матрица U, осуществляющая преобразование подобия матрицы A к матрице Λ состоит из нормированных собственных векторов матрицы A и имеет вид

$$U = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} \\ \frac{i}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}.$$

Непосредственно проверяется, что $A = U\Lambda U^*$.

Унитарная (ортогональная) диагонализуемость влечет за собой диагонализуемость, но не наоборот.

Пример 5.8 (Диагонализуемость). Матрица

$$A = \begin{bmatrix} 2 & 0 \\ 4-2i & i \end{bmatrix}$$

имеет минимальный полином

$$m_A(\lambda) = (2-\lambda)(i-\lambda),$$

который разлагается на линейные множители. Из этого следует, что матрица A диагонализуема. Преобразование подобия,

осуществляющее диагонализацию матрицы A, выполняется с помощью невырожденной матрицы $S = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}$:

$$S^{-1}AS = \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 4-2i & i \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & i \end{bmatrix} = \Lambda.$$

Матрица A не является нормальной, так как

$$AA^* = \begin{bmatrix} 4 & 8+4i \\ 8-4i & 21 \end{bmatrix} \neq \begin{bmatrix} 24 & -2+4i \\ -2-4i & 1 \end{bmatrix} = A^*A.$$

Значит, согласно спектральной теореме для нормальных матриц, матрица A не является унитарно диагонализуемой.

Определение 5.7. *Сингулярные числа* $\sigma_i,\ i = \overline{1, \min\{n,m\}}$ прямоугольной матрицы $A \in \mathbb{C}^{n\times m}$ — это неотрицательные квадратные корни из собственных значений матрицы A^*A или AA^*:

$$\sigma_i = \sqrt{\lambda_i(A^*A)},\ i = \overline{1,n}.$$

Теорема 5.8. *Сингулярные числа нормальной матрицы равны по модулю соответствующим характеристическим числам:*

$$\sqrt{\lambda_i\left(A^*A\right)} = \left|\lambda_i\left(A\right)\right|.$$

Доказательство. A — нормальна, значит, унитарно диагонализуема: $A = U^*\Lambda U$, где $\Lambda = \operatorname{diag}\left\{\lambda_1\left(A\right), \ldots, \lambda_n\left(A\right)\right\}$, U — унитарная матрица. Тогда

$$\begin{aligned} \lambda_i\left(A^*A\right) &= \lambda_i\left(\left(U^*\Lambda U\right)^* U^*\Lambda U\right) = \lambda_i\left(U^*\Lambda^*\Lambda U\right) = \\ &= \lambda_i\left(U^*\operatorname{diag}\left\{\left|\lambda_1\right|^2, \ldots, \left|\lambda_n\right|^2\right\} U\right) = \left|\lambda_i\right|^2. \end{aligned}$$ □

Задачи и упражнения

1. Доказать, что для любой ортогональной системы $\{y_1, \ldots, y_k\}$ ненулевых векторов система $\{x_1, \ldots, x_k\}$, где $x_i = \langle y_i^* y_i \rangle^{-\frac{1}{2}} y_i$, $i = \overline{1, k}$, будет ортонормированной.
2. Доказать, что любая ортогональная система ненулевых векторов линейно независима.
3. Доказать, что для любой ортогональной системы $x_1, \ldots, x_k \in \mathbb{C}^n$ либо $k \le n$, либо по меньшей мере $k - n$ векторов x_i равны нулю.
4. Доказать, что если система векторов $\{x_1, \ldots, x_k\}$ ортонормированная и матрица $U \in \mathbb{C}^{n\times n}$ — унитарна, то система $\{Ux_1, \ldots, Ux_k\}$ также ортонормирована.
5. Выполнить QR-разложение следующих матриц:

   a) $\begin{bmatrix} 3 & 1 \\ -2 & 0 \end{bmatrix}$, b) $\begin{bmatrix} -2 & 1 & 2 \\ -1 & -2 & -1 \\ 1 & 1 & 1 \end{bmatrix}$, c) $\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & -2 \\ 1 & 3 & -1 \end{bmatrix}$.
6. Привести какой-либо пример вещественной матрицы $A \in \mathbb{R}^{2\times 2}$, которая не является нормальной. Привести также пример вещественной матрицы $B \in \mathbb{R}^{2\times 2}$, которая является нормальной, не будучи симметричной, *кососимметричной* (т.е. такой, что $A^T = -A$) или ортогональной.

7*. Доказать следующие утверждения:

   a) матрица $A \in \mathbb{C}^{n\times n}$ нормальна в том и только том случае, когда векторы Ax и A^*x имеют одинаковую длину для любого $x \in \mathbb{C}^n$;

   b) матрица $A \in \mathbb{C}^{n\times n}$ нормальна в том и только в том случае, когда $(Ax)^*(Ay) = (A^*x)^*(A^*y)$ для всех $x, y \in \mathbb{C}^n$. Что это означает геометрически? Как связано с a)?

8. Привести пример матриц, которые являются диагонализуемыми, но не являются унитарно диагонализуемыми.
9. Доказать, что если матрица $A \in \mathbb{C}^{n\times n}$ нормальна, то равенство $Ax = 0$ равносильно $A^*x = 0$. Это означает, что нуль-пространства матриц A и A^* совпадают. Рассмотреть

   $$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

   и показать, что это не верно в общем случае.

10. Доказать, что две нормальные матрицы подобны тогда и только тогда, когда они унитарно подобны.
11. Доказать, что матрица $A \in \mathbb{C}^{n\times n}$ нормальна в том и только том случае, когда всякий собственный вектор для A является также собственным вектором для A^*.

12*. Используя спектральную теорему для нормальных матриц, доказать, что если матрица $A \in \mathbb{C}^{n\times n}$ нормальна и $Ax = \lambda x$, $Ay = \mu y$ при $\lambda \ne \mu$, то x и y ортогональны.

13. Доказать, что следующие матрицы унитарно диагонализуемы, и построить для них спектральное разложение (5.7):

    a) $A = \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix}$; b) $C = \begin{bmatrix} \frac{1}{4} & \frac{\sqrt{3}}{4} \\ \frac{\sqrt{3}}{4} & \frac{3}{4} \end{bmatrix}$; c) матрица B из упражнения 6.
14. Доказать, что матрица $A \in \mathbb{C}^{n\times n}$ унитарна тогда и только тогда, когда ее можно записать в виде $A = e^{iH}$, где H — эрмитова ($H^* = H$).
15. Доказать следующую теорему

    Теорема 5.9* (Вещественный вариант теоремы Шура). *Для любой матрицы $A \in \mathbb{R}^{n\times n}$ существует ортогональная матрица $Q \in \mathbb{R}^{n\times n}$, такая, что*

    $$Q^T A Q = \begin{bmatrix} A_1 & * & \ldots & * \\ 0 & A_2 & \ldots & * \\ 0 & 0 & \ldots & * \\ 0 & 0 & \ldots & A_k \end{bmatrix},$$

    где для каждого i матрица A_i имеет размер 1×1 или 2×2, отвечая соответственно вещественному собственному значению или невещественной паре комплексно сопряженных собственных значений матрицы A. Блоки A_i можно расположить в любом заданном порядке.

16. Доказать, что для того, чтобы матрица $A \in \mathbb{C}^{n\times n}$ была нормальной, необходимо и достаточно, чтобы любая унитарно подобная ей матрица была нормальной. Т.е., унитарное подобие не выводит из класса нормальных матриц.

Вопросы для повторения

1. Какие векторы называются ортогональными? Ортонормированными?
2. Каким образом произвольную линейно независимую систему векторов можно преобразовать к ортонормированной системе, имеющей ту же линейную оболочку, что и исходная система? Опишите процесс преобразования.
3. Чем характеризуются собственные значения унитарной матрицы? Определитель унитарной матрицы?
4. Пусть U, $V \in \mathbb{C}^{n\times n}$— унитарные матрицы. Какому классу матриц принадлежат U^*, U^{-1}, UV?
5. Приведите примеры матрицы перестановок.
6. Когда матрица $A \in \mathbb{C}^{n\times n}$ унитарно подобна матрице $B \in \mathbb{C}^{n\times n}$?
7. Когда матрица $A \in \mathbb{C}^{n\times n}$ ортогонально подобна матрице $B \in \mathbb{C}^{n\times n}$?
8. Сформулируйте следствия из теоремы Шура (о «почти диагонализуемости» произвольной матрицы).
9. Матрицы из каких известных вам классов являются нормальными?
10. Какие из следующих соотношений верны для произвольной унитарной матрицы U:
    a) $U^*U = UU^*$; b) $U^* = U^T$; c) $U^* = \bar{U}$; d) $U^* = U^{-1}$;
    e) $U = U^*$; f) $|\det U| = 1$; g) $|\operatorname{tr} U| = 1$; i) $|\lambda_i(U)| = 1$;
    j) $|\sigma_i| = 1$?
11. Какие из следующих утверждений справедливы для любых унитарных матриц $U, V \in \mathbb{C}^{n\times n}$:
    a) $UV = E_{n\times n}$; b) UV — унитарна; c) $U^*V = E_{n\times n}$;
    d) $UV^* = E_{n\times n}$; e) U^*V — унитарна; f) U^* — унитарна?

Необходимо усвоить.

Основные определения и свойства матриц

- нормальная
- ортогональная
- перестановок
- унитарная
- унитарно диагонализуемая, ортогонально диагонализуемая
- матрицы ортогонально подобные
- матрицы унитарно подобные
- подобие унитарное

Формулировку и трактовку теорем

- о линейной независимости ортонормированных векторов
- о критериях унитарности матрицы
- теорема Шура об унитарной триангуляризации
- о QR-разложении
- спектральная теорема для нормальных матриц

Основные умения и навыки

- определять принадлежность заданной матрицы определенному классу (унитарная, ортогональная, нормальная)
- распознавать по матрице возможность построения для нее того или иного разложения и строить (если возможно) это разложение (спектральное разложение, унитарную триангуляризацию, QR-разложение).

Глава 6

Эрмитовы и симметричные матрицы

Специфические свойства эрмитовых матриц делают удобным их применение во многочисленных приложениях. Так, например, связанное с эрмитовыми матрицами сингулярное разложение используется при решении систем уравнений, приближении методом наименьших квадратов, сжатии изображений, в поисковых системах. Сингулярное разложение позволяет вычислять обратные и псевдообратные матрицы большого размера, что делает его полезным инструментом при решении задач регрессионного анализа в статистике.

Рекомендуемая литература:
[1]*, [3]*, [5], [7], [9], [10], [11], [13], [15]

Цель изучения

Освоить понятие и основные свойства эрмитовых и симметричных матриц, а также соответствующих линейных преобразований, получить навыки в построении связанных с эрмитовыми матрицами матричных разложений.

Следует повторить

- свойства действительных и комплексных чисел
- изометричность унитарных матриц
- понятие и свойства нормальных матриц
- спектральная теорема для нормальных матриц

6.1. Определение и свойства эрмитовых матриц

Определение 6.1. Матрица $A \in \mathbb{C}^{n\times n}$ называется *эрмитовой*, если $A = A^*$, и *косоэрмитовой*, если $A = -A^*$.

Свойства эрмитовых матриц

1^0. Матрицы $A + A^*$, AA^*, A^*A эрмитовы для любой матрицы $A \in \mathbb{C}^{n\times n}$.

2^0. Если матрица A эрмитова, то ее степень A^k эрмитова для всех $k = 1, 2, \ldots$. Если A также невырожденная, то A^{-1} тоже эрмитова.

3^0. Если A и B — эрмитовы, то матрица $\alpha A + \beta B$ эрмитова для любых вещественных α, β.

4^0. Матрица $A - A^*$ косоэрмитова для любой матрицы $A \in \mathbb{C}^{n\times n}$.

5^0. Если A и B — косоэрмитовы, то матрица $\alpha A + \beta B$ косоэрмитова для любых вещественных α, β.

6^0. Если A — эрмитова, то iA — косоэрмитова.

7^0. Если A — косоэрмитова, то iA — эрмитова.

8^0. Любую матрицу $A \in \mathbb{C}^{n\times n}$ можно записать в виде:

$$A = \frac{1}{2}(A + A^*) + \frac{1}{2}(A - A^*) \equiv H(A) + S(A),$$

где $H(A) = \frac{1}{2}(A + A^*)$ — эрмитова часть матрицы A, $S(A) = \frac{1}{2}(A - A^*)$ — косоэрмитова часть матрицы A.

9^0. Если A — эрмитова, то все элементы на ее главной диагонали вещественны. Для того чтобы задать все n^2 элементов матрицы A, достаточно указать n вещественных чисел (диагональные элементы) и $\frac{n(n-1)}{2}$ комплексных чисел (внедиагональные).

Упражнение 6.1. Доказать свойства эрмитовых матриц непосредственно по определению.

Пример 6.1 (Эрмитова и косоэрмитовы матрицы). Матрица

$$\begin{bmatrix} 1 & -i & -4 \\ i & -2 & 1+2i \\ -4 & 1-2i & 3 \end{bmatrix}$$

является эрмитовой. Матрица

$$\begin{bmatrix} -i & -1+i & -4 \\ 1+i & -2i & 1-2i \\ 4 & -1-2i & 0 \end{bmatrix}$$

является косоэрмитовой.

Теорема 6.1 (Об эрмитовом разложении). *Любую матрицу $A \in \mathbb{C}^{n\times n}$ можно записать единственным образом в виде $A = H + iT$, где обе матрицы H и T эрмитовы. Имеется также единственное представление вида $A = B + C$, в котором матрица B — эрмитова, а матрица C — косоэрмитова.*

Доказательство. Существование представления следует из свойства 8^0, если положить $B = H(A)$, $C = S(A)$, свойства 6^0 и равенства $C = -iiC$ (поскольку iC — эрмитова), а также свойства 7^0: $H = H(A)$, $T = -iC$.

Докажем единственность. Если $A = K + iF$ — еще одно представление, где K и F — эрмитовы, то

$$\begin{aligned} 2H = A + A^* &= (K + iF) + (K + iF)^* = \\ &= K + iF + K^* - iF^* = 2K, \end{aligned}$$

откуда следует, что

$$H = S.$$

Аналогично устанавливается равенство $F = T$ и существование единственного представления $A = B + C$. □

Упражнение 6.2. Доказать, что если $A, B \in \mathbb{C}^{n\times n}$ — эрмитовы, то AB — эрмитова тогда и только тогда, когда A и B коммутируют (перестановочны).

Пример 6.2 (Эрмитово разложение). Запишем эрмитово разложение матрицы

$$A = \begin{bmatrix} 2 & i & 0 \\ 3-i & 2i & -4 \\ 3 & 0 & 1 \end{bmatrix}.$$

Используя свойство 8^0, представим матрицу A в виде суммы эрмитовой и косоэрмитовой матрицы $A = H(A) + S(A)$, где

$$\begin{aligned} H(A) &= \frac{1}{2}(A + A^*) = \\ &= \frac{1}{2}\left(\begin{bmatrix} 2 & i & 0 \\ 3-i & 2i & -4 \\ 3 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 2 & 3+i & 3 \\ -i & -2i & 0 \\ 0 & -4 & 1 \end{bmatrix}\right) = \\ &= \begin{bmatrix} 2 & \frac{3}{2}+i & \frac{3}{2} \\ \frac{3}{2}-i & 0 & -2 \\ \frac{3}{2} & -2 & 1 \end{bmatrix}, \end{aligned}$$

$$S(A) = \frac{1}{2}(A - A^*) =$$
$$= \frac{1}{2}\left(\begin{bmatrix} 2 & i & 0 \\ 3-i & 2i & -4 \\ 3 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 2 & 3+i & 3 \\ -i & -2i & 0 \\ 0 & -4 & 1 \end{bmatrix}\right) =$$
$$= \begin{bmatrix} 0 & -\frac{3}{2} & -\frac{3}{2} \\ \frac{3}{2} & 2i & -2 \\ \frac{3}{2} & 2 & 0 \end{bmatrix}.$$

Так как матрица $S(A)$ — косоэрмитова, то согласно свойству 7^0, матрица

$$T(A) = -iS(A) = \frac{S(A)}{i} = \begin{bmatrix} 0 & \frac{3}{2}i & \frac{3}{2}i \\ -\frac{3}{2}i & 2 & 2i \\ -\frac{3}{2}i & -2i & 0 \end{bmatrix}$$

— эрмитова.

Таким образом, для матрицы A имеем представление

$$A = H(A) + S(A) = H(A) + i\frac{S(A)}{i} = H(A) + iT(A).$$

Теорема 6.2 (Критерии эрмитовости). *Матрица $A \in \mathbb{C}^{n\times n}$ эрмитова тогда и только тогда, когда выполнено хотя бы одно из следующих условий:*

a) функция x^*Ax принимает вещественные значения для любого $x \in \mathbb{C}^n$;

b) матрица A нормальна и все ее собственные значения вещественны;

c) матрица S^*AS эрмитова для любой матрицы $S \in \mathbb{C}^{n\times n}$.

Доказательство. Необходимость.

a) Рассмотрим число, комплексно-сопряженное к числу x^*Ax:

$$\overline{x^*Ax} = (x^*Ax)^* = x^*A^*x \overset{A^*=A}{=\!=\!=\!=} x^*Ax.$$

Таким образом, число x^*Ax совпадает с комплексно сопряженным числом $\overline{x^*Ax}$ и, следовательно, является вещественным.

b) Пусть λ — собственное значение эрмитовой матрицы A, x — соответствующий ему нормированный собственный вектор:

$$Ax = \lambda x, \ \ x^*x = 1.$$

Умножим второе равенство на λ:

$$\lambda = \lambda x^*x = \ x^*\lambda x = x^*Ax \in \mathbb{R}$$

(в силу доказанного утверждения а).

Нормальность матрицы А вытекает из справедливости равенства: $AA^* = (A = A^*, A^* = A) = A^*A$.

c) Для произвольной матрицы $S \in \mathbb{C}^{n\times n}$ справедливо:

$$(S^*AS)^* = S^*A^*S \overset{A^*=A}{=\!=\!=\!=} S^*AS,$$

следовательно S^*AS— эрмитова.

Достаточность.

Рассмотрим утверждение а) для вектора $(x+y) \in \mathbb{C}^n$:

$$(x+y)^*A(x+y) =$$
$$= (x^*Ax + y^*Ay) + (x^*Ay + y^*Ax) \in \mathbb{R} \ \ \forall\, x, y \in \mathbb{C}^n.$$

Поскольку $x^*Ax \in \mathbb{R}, \ y^*Ay \in \mathbb{R}$ по предположению, то

$$x^*Ay + y^*Ax \in \mathbb{R}, \ \forall x, y \in \mathbb{C}^n. \tag{6.1}$$

Возьмем в (6.1) $x = e_k, \ y = e_j$ (e_k — k-й вектор естественного базиса $\mathbb{C}^n$):

$$e_k^*Ae_j + e_j^*Ae_k = a_{kj} + a_{jk} \in \mathbb{R} \Rightarrow \operatorname{Im} a_{kj} = -\operatorname{Im} a_{jk}. \tag{6.2}$$

Возьмем $x = ie_k, \ y = e_j$:

$$-ie_k^*Ae_j + ie_j^*Ae_k = -ia_{kj} + ia_{jk} \in \mathbb{R} \Rightarrow \operatorname{Re} a_{kj} = \operatorname{Re}\, a_{jk}. \tag{6.3}$$

Из (6.2) и (6.3) следует, что $a_{kj} = \overline{a_{jk}}$, $\forall j, k = \overline{1, n}$, а, следовательно, $A = A^*$.

b) По спектральной теореме для нормальных матриц нормальная матрица унитарно диагонализуема, т.е. $A = U\Lambda U^*$, где $U \in \mathbb{C}^{n\times n}$ — унитарная матрица, $\Lambda = \operatorname{diag}\{\lambda_1, .., \lambda_n\} \in \mathbb{R}^{n\times n}$, $\lambda_i \in \sigma(A)$.

Рассмотрим $A^* = (U\Lambda U^*)^* = U\Lambda^* U^* \overset{\Lambda \in \mathbb{R}^{n\times n}}{=\!=\!=\!=} U\Lambda U^* = A$, следовательно А — эрмитова.

c) Положим $S = E_n$: $S^*AS = (S^*AS)^* \Rightarrow A = A^*$. □

Теорема 6.3 (Спектральное разложение эрмитовых матриц). *Матрица $A \in \mathbb{C}^{n\times n}$ эрмитова тогда и только тогда, когда существует унитарная матрица $U \in \mathbb{C}^{n\times n}$ и вещественная диагональная матрица $\Lambda \in \mathbb{R}^{n\times n}$ такие, что*

$$A = U\Lambda U^*. \tag{6.4}$$

Доказательство. Необходимость: A — эрмитова, следовательно, A — нормальная и $\sigma(A) \subset \mathbb{R}$. По спектральной теореме для нормальных матриц $A = U\Lambda U^*$, $\Lambda \in \mathbb{R}^{n\times n}$.

Достаточность.

$$A^* = (U\Lambda U^*)^* = U\Lambda^* U^* \overset{\Lambda \in \mathbb{R}^{n\times n}}{=\!=\!=\!=} U\Lambda U^* = A,$$

следовательно, A — эрмитова. □

Следствие 6.1. *Матрица A вещественна и эрмитова (симметрична) тогда и только тогда, когда существуют ортогональная матрица $P \in \mathbb{R}^{n\times n}$ и вещественная диагональная матрица $\Lambda \in \mathbb{R}^{n\times n}$ такие, что*

$$A = P\Lambda P^T.$$

Упражнение 6.3. Доказать следствие 6.1.

В силу утверждения b) теоремы 6.2 к эрмитовой матрице применимы все результаты раздела 5.3.

Пример 6.3 (Спектральное разложение эрмитовой матрицы). Эрмитова матрица

$$A = \begin{bmatrix} -1 & -3i \\ 3i & -1 \end{bmatrix}$$

имеет собственные значения $\lambda_1 = 2$, $\lambda_2 = -4$. Значит, она унитарно подобна вещественной диагональной матрице

$$\Lambda = \begin{bmatrix} 2 & 0 \\ 0 & -4 \end{bmatrix}.$$

Для построения спектрального разложения найдем собственные векторы.

Для $\lambda_1 = 2$ имеем:

$$(A - \lambda_1 E) = \begin{bmatrix} -3 & -3i \\ 3i & -3 \end{bmatrix} \sim \begin{bmatrix} i & -1 \end{bmatrix}.$$

Первый собственный вектор

$$x^{(1)=} \begin{bmatrix} \alpha \\ i\alpha \end{bmatrix}.$$

Нормируем этот вектор:

$$\|x^{(1)}\| = \sqrt{\alpha^2 + \alpha^2} = \alpha\sqrt{2} = 1.$$

Значит, вектор будет нормированным при $\alpha = \frac{1}{\sqrt{2}}$:

$$x^{(1)=} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{i}{\sqrt{2}} \end{bmatrix}.$$

Аналогично находим второй нормированный собственный вектор для $\lambda_2 = -4$:

$$(A - \lambda_2 E) = \begin{bmatrix} 3 & -3i \\ 3i & 3 \end{bmatrix} \sim \begin{bmatrix} i & 1 \end{bmatrix},$$

$$x^{(2)}= \begin{bmatrix} \alpha \\ -i\alpha \end{bmatrix}, \quad \|x^{(2)}\| = \alpha\sqrt{2} = 1.$$

$$x^{(2)}= \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -i\frac{1}{\sqrt{2}} \end{bmatrix}.$$

Заметим, что

$$x^{(1)*}x^{(2)} = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{i}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{i}{\sqrt{2}} \end{bmatrix} = \frac{1}{2} - \frac{1}{2} = 0,$$

т.е. собственные векторы ортогональны.

Составим из них унитарную матрицу

$$U = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{i}{\sqrt{2}} & \frac{-i}{\sqrt{2}} \end{bmatrix}.$$

Итак, спектральное разложение матрицы A имеет вид:

$$A = U\Lambda U^{*}U = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{i}{\sqrt{2}} & \frac{-i}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & -4 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{-i}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{-i}{\sqrt{2}} \end{bmatrix}.$$

6.2. Вариационные описания собственных значений эрмитовых матриц и их приложения

Пусть $A \in \mathbb{C}^{n\times n}$ — эрмитова матрица, $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{R}$ — ее собственные значения. Согласно утверждению b) теоремы 6.2 их можно считать упорядоченными:

$$\lambda_{\min} = \lambda_1 \le \lambda_2 \le \ldots \le \lambda_n = \lambda_{\max}. \tag{6.5}$$

Теорема 6.4 (Рэлея—Ритца). *Пусть матрица A — эрмитова и ее собственные значения упорядочены как в (6.5). Тогда*

$$\lambda_1 x^{*}x \le x^{*}Ax \le \lambda_n x^{*}x, \ \forall x \in \mathbb{C}^n, \tag{6.6}$$

$$\lambda_{\max} = \lambda_n = \max_{x\neq 0} \frac{x^{*}Ax}{x^{*}x} = \max_{x^{*}x=1} x^{*}Ax, \tag{6.7}$$

$$\lambda_{\min} = \lambda_1 = \min_{x\neq 0} \frac{x^{*}Ax}{x^{*}x} = \min_{x^{*}x=1} x^{*}Ax. \tag{6.8}$$

Доказательство. Поскольку матрица A — эрмитова, то по спектральной теореме для эрмитовых матриц существуют такие унитарная матрица $U \in \mathbb{C}^{n\times n}$ и диагональная матрица $\Lambda = \operatorname{diag}\{\lambda_1, \ldots, \lambda_n\}$, что $A = U\Lambda U^{*}$. При любом векторе $x \in \mathbb{C}^n$ верны равенства

$$x^{*}Ax = x^{*}U\Lambda U^{*}x = (U^{*}x)^{*}\,\Lambda\,(U^{*}x) = \sum_{i=1}^{n} \lambda_i \left|(U^{*}x)_i\right|^2.$$

Каждый сомножитель $\left|(U^{*}x)_i\right|^2$ неотрицателен, поэтому

$$\lambda_{\min}\sum_{i=1}^{n} \left|(U^{*}x)_i\right|^2 \le x^{*}Ax =$$

$$= \sum_{i=1}^{n} \lambda_i \left|(U^{*}x)_i\right|^2 \le \lambda_{\max}\sum_{i=1}^{n} \left|(U^{*}x)_i\right|^2.$$

Поскольку матрица U унитарна, а унитарные матрицы изометричны (теорема 5.2), то имеют место равенства

$$\sum_{i=1}^{n} \left|(U^{*}x)_i\right|^2 = \sum_{i=1}^{n} |x_i|^2 = x^{*}x.$$

Таким образом, справедливы соотношения:

$$\lambda_1 x^{*}x = \lambda_{\min}x^{*}x \ \le \ x^{*}Ax \le \ \lambda_{\max}x^{*}x = \lambda_n x^{*}x, \tag{6.9}$$

что и доказывает (6.6).

Покажем, что оценки в соотношении (6.6) точны.

Действительно, если x — собственный вектор матрицы A, соответствующий собственному значению λ_1, то

$$x^*Ax = x^*\lambda_1 x = \lambda_1 x^*x.$$

Точность оценки сверху устанавливается аналогично.

Остальные утверждения являются простыми следствиями соотношений (6.6).

Покажем, например, справедливость (6.7). При $x \neq 0$ из правой части соотношения (6.6) имеем:

$$\frac{x^*Ax}{x^*x} \le \lambda_n.$$

Это неравенство обращается в равенство, когда x — собственный вектор матрицы A, соответствующий собственному значению λ_n. Следовательно,

$$\max_{x\neq 0} \frac{x^*Ax}{x^*x} = \lambda_n. \tag{6.10}$$

Наконец, при $x \neq 0$ можно перейти к нормированному вектору:

$$\frac{x^*Ax}{x^*x} = \left(\frac{x}{\sqrt{x^*x}}\right)^* A \left(\frac{x}{\sqrt{x^*x}}\right), \quad \left(\frac{x}{\sqrt{x^*x}}\right)^* \left(\frac{x}{\sqrt{x^*x}}\right) = 1.$$

Таким образом, равенство (6.10) эквивалентно следующему

$$\max_{x^*x=1} x^*Ax = \lambda_n. \tag{6.11}$$

Рассуждения для минимального собственного значения λ_1 аналогичны. □

Геометрическая интерпретация равенства (6.11) состоит в том, что число λ_n есть наибольшее значение функции x^*Ax, когда вектор x пробегает единичную сферу в пространстве $\mathbb{C}^n$ (которая является компактным множеством).

Пример 6.4* (Собственные значения эрмитовой матрицы). Продемонстрируем, как можно применить теорему Рэлея—Ритца для нахождения минимального и максимального собственных значений эрмитовой матрицы A.

Рассмотрим эрмитову матрицу

$$A = \begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{bmatrix}.$$

Отношение Рэлея—Ритца $\frac{x^*Ax}{x^*x}$, $x^*x = 1$, для этой матрицы имеет вид:

$$\frac{x^*Ax}{x^*x} = x^TAx = [x_1, x_2] \begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} =$$
$$= \frac{1}{2}x_1^2 - \sqrt{3}x_1x_2 - \frac{1}{2}x_2^2.$$

Согласно теореме Рэлея—Ритца для нахождения $\lambda_{\min}(A)$ и $\lambda_{\max}(A)$ надо найти минимальное и максимальное значения функции

$$f(x_1, x_2) = \frac{1}{2}x_1^2 - \sqrt{3}x_1x_2 - \frac{1}{2}x_2^2$$

при условии, что x_1 и x_2 связаны соотношением

$$x^*x = x_1^2 + x_2^2 = 1.$$

Для нахождения условных экстремумов функции $f(x_1, x_2)$ выразим из последнего равенства x_1 и подставим его в выражение для $f(x_1, x_2)$. Имеем:

$$x_1 = \pm\sqrt{1 - x_2^2}, \; x_2 \in [-1, 1], \tag{6.12}$$

$$f(x_1, x_2) = \frac{1}{2} - x_2^2 \mp x_2\sqrt{3\left(1 - x_2^2\right)} = \bar{f}(x_2), \; x_2 \in [-1, 1].$$

Используя необходимые условия экстремума первого порядка для функции $\bar{f}(x_2)$:

$$\bar{f}'(x_2) = -2x_2 \mp \sqrt{3}\frac{1-2x_2^2}{\sqrt{1-x_2^2}} = 0,$$

с учетом выражения (6.12) имеем четыре подозрительные на экстремум точки:

$$a = \begin{bmatrix} \frac{1}{2} \\ \frac{\sqrt{3}}{2} \end{bmatrix},\ b = \begin{bmatrix} -\frac{1}{2} \\ -\frac{\sqrt{3}}{2} \end{bmatrix},\ c = \begin{bmatrix} \frac{\sqrt{3}}{2} \\ -\frac{1}{2} \end{bmatrix},\ d = \begin{bmatrix} -\frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{bmatrix}.$$

Кроме них подозрительными являются точки, у которых координата x_2 является границей отрезка $[-1,1]$, т.е. $(0,-1)$ и $(0,1)$. Вычислим значения функции $f(x_1,x_2)$ в подозрительных точках:

$$f(a) = f(b) = -1,\ f(c) = f(d) = 1,\ f(0,-1) = f(0,1) = -\frac{1}{2}.$$

Таким образом,

$$\min_{x_1^2+x_2^2=1} f(x_1,x_2) = -1,\ \max_{x_1^2+x_2^2=1} f(x_1,x_2) = 1.$$

Непосредственное вычисление показывает, что матрица A имеет два различных собственных значения $\lambda_1 = -1,\ \lambda_2 = 1$, что совпадает со значениями, найденными по теореме Рэлея—Ритца.

Теорема 6.5* (Куранта—Фишера). *[10, стр. 215] Для собственного значения λ_k $(k = \overline{1,n})$ эрмитовой матрицы $A \in \mathbb{C}^{n\times n}$ имеют место представления*

$$\lambda_k = \min_{\omega_1,\dots,\omega_{n-k}\in\mathbb{C}^n} \max_{\substack{x\neq 0,\, x\in\mathbb{C}^n \\ x\perp\omega_1,\dots,\omega_{n-k}}} \frac{x^*Ax}{x^*x}, \tag{6.13}$$

$$\lambda_k = \max_{\omega_1,\dots,\omega_{k-1}\in\mathbb{C}^n} \min_{\substack{x\neq 0,\, x\in\mathbb{C}^n \\ x\perp\omega_1,\dots,\omega_{k-1}}} \frac{x^*Ax}{x^*x}. \tag{6.14}$$

Замечание. При $k = n$ в (6.13) и при $k = 1$ в (6.14) имеем равенства (6.7), (6.8) из теоремы Релея—Ритца.

Среди многочисленных важных приложений теоремы Куранта—Фишера одно из простейших связано с задачей сравнения собственных значений матриц $A+B$ и A.

Для произвольной матрицы A обозначим ее собственные значения через $\lambda_i(A)$.

Теорема 6.6 (Вейля). *Пусть матрицы $A, B \in \mathbb{C}^{n\times n}$ эрмитовы и собственные значения $\lambda_i(A)$, $\lambda_i(B)$, $\lambda_i(A+B)$ упорядочены по возрастанию как в (6.2). Тогда при каждом $k = 1,2,\dots,n$ выполняются неравенства*

$$\lambda_k(A) + \lambda_1(B) \le \lambda_k(A+B) \le \lambda_k(A) + \lambda_n(B). \tag{6.15}$$

Доказательство. Для любого ненулевого вектора $x \in \mathbb{C}^n$ из отношений (6.7), (6.8) Рэлея—Ритца для эрмитовой матрицы B следует:

$$\lambda_1(B) \le \frac{x^*Bx}{x^*x} \le \lambda_n(B). \tag{6.16}$$

Следовательно, при любом $k = 1,2,\dots,n$

$$\lambda_k(A+B) \overset{(6.13)}{=} \min_{\omega_1,\dots,\omega_{n-k}\in\mathbb{C}^n} \max_{\substack{x\neq 0,\, x\in\mathbb{C}^n \\ x\perp\omega_1,\dots,\omega_{n-k}}} \frac{x^*(A+B)x}{x^*x} =$$

$$= \min_{\omega_1,\dots,\omega_{n-k}\in\mathbb{C}^n} \max_{\substack{x\neq 0,\, x\in\mathbb{C}^n \\ x\perp\omega_1,\dots,\omega_{n-k}}} \left(\frac{x^*Ax}{x^*x} + \frac{x^*Bx}{x^*x}\right) \overset{(6.16)}{\ge}$$

$$\ge \min_{\omega_1,\dots,\omega_{n-k}\in\mathbb{C}^n} \max_{\substack{x\neq 0,\, x\in\mathbb{C}^n \\ x\perp\omega_1,\dots,\omega_{n-k}}} \left(\frac{x^*Ax}{x^*x} + \lambda_1(B)\right) =$$

$$= \min_{\omega_1,\dots,\omega_{n-k}\in\mathbb{C}^n} \max_{\substack{x\neq 0,\, x\in\mathbb{C}^n \\ x\perp\omega_1,\dots,\omega_{n-k}}} \frac{x^*Ax}{x^*x} + \lambda_1(B) =$$

$$\overset{(6.13)}{=} \lambda_k(A) + \lambda_1(B).$$

Оценка сверху доказывается аналогично. □

Пример 6.5 (Точность оценок теоремы Вейля). Покажем, что левое неравенство в (6.15) может обращаться в равенство.

Пусть $\{ u_1, \ u_2, \ldots, u_n \}$ — ортонормированная система собственных векторов матрицы A и $Au_k = \lambda_k(A)u_k$. Существование такой системы гарантируется спектральной теоремой для нормальных матриц (теорема 5.7), поскольку матрица A — эрмитова, а значит, нормальная.

Рассмотрим эрмитову матрицу $B = \alpha u_k u_k^*$. Эта матрица имеет собственные значения

$$\lambda_j(B) = 0, \ j = \overline{1,n}, \ j \neq k, \quad \lambda_k(B) = \alpha,$$

которые соответствуют собственным векторам u_j, $j = \overline{1,n}$. Действительно, в силу ортонормированности системы $\{ u_1, \ u_2, \ldots, u_n \}$, имеем:

$$\begin{aligned} Bu_j &= \alpha u_k u_k^* u_j = 0 = \lambda_j(B)u_j, \ j = \overline{1,n}, \ j \neq k, \\ Bu_k &= \alpha u_k u_k^* u_k = \alpha u_k. \end{aligned}$$

Тогда при $\alpha < 0$ имеем $\lambda_{\min}(B) = \lambda_1(B) = \alpha$ и левая часть неравенства (6.16) принимает значение $\lambda_k(A) + \alpha$.

Найдем теперь $\lambda_k(A + B)$. Используя первое соотношение теоремы Куранта—Фишера, получаем:

$$\begin{aligned} \lambda_k(A+B) &= \min_{\omega_1,\ldots,\omega_{n-k}\in\mathbb{C}^n} \ \max_{\substack{x\neq 0,\, x\in\mathbb{C}^n \\ x\perp\omega_1,\ldots,\omega_{n-k}}} \frac{x^*(A+B)x}{x^*x} = \\ &= \min_{\omega_1,\ldots,\omega_{n-k}\in\mathbb{C}^n} \ \max_{\substack{x\neq 0,\, x\in\mathbb{C}^n \\ x\perp\omega_1,\ldots,\omega_{n-k}}} \frac{x^*Ax}{x^*x} + \\ &+ \min_{\omega_1,\ldots,\omega_{n-k}\in\mathbb{C}^n} \ \max_{\substack{x\neq 0,\, x\in\mathbb{C}^n \\ x\perp\omega_1,\ldots,\omega_{n-k}}} \frac{x^*Bx}{x^*x} = \\ &= \lambda_k(A) + \lambda_k(B) = \lambda_k(A) + \alpha, \end{aligned}$$

что совпадает с левой частью неравенства (6.15).

Аналогично, рассматривая $B = \alpha u_k u_k^*$ для $\alpha > 0$, можно показать, что правое неравенство в (6.15) может обращаться в равенство.

6.3*. Комплексные симметричные матрицы

Определение 6.2. Матрица $A \in \mathbb{C}^{n\times n}$ называется *симметричной,* если $A = A^T$.

Пример 6.6 (Симметричная матрица вторых производных). Пусть $f : D \to \mathbb{R}$ — дважды непрерывно дифференцируемая функция в некоторой области $D \subseteqq \mathbb{R}^n$. Вещественную функциональную матрицу

$$H(x) = (h_{ij}(x)) \equiv \left[\frac{\partial^2 f(x)}{\partial x_i \partial x_j}\right] \in \mathbb{R}^{n\times n}$$

называют *гессианом* (или *матрицей Гессе*) этой функции. В силу равенства смешанных производных:

$$\frac{\partial^2 f(x)}{\partial x_i \partial x_j} = \frac{\partial^2 f(x)}{\partial x_j \partial x_i}, \ \forall i, j = 1, 2, \ldots, n,$$

для всех индексов $i, j = 1, 2, \ldots, n$ верно равенство

$$h_{ij}(x) = h_{ji}(x), \forall x \in D.$$

Таким образом, гессиан является симметричной матрицей для любого $x \in D$.

Пример 6.7 (Симметричная матрица квадратичной формы). Рассмотрим некоторую матрицу $A \in \mathbb{C}^{n \times n}$ и связанную с ней квадратичную форму на $\mathbb{C}$:

$$Q(x) \equiv x^T A x = \sum_{i,j=1}^{n} a_{ij} x_i x_j =$$

$$= \sum_{i,j=1}^{n} \frac{1}{2}(a_{ij} + a_{ji}) x_i x_j = x^T \left[\frac{1}{2}(A + A^T) \right] x.$$

Таким образом, матрицы A и $\frac{1}{2}(A + A^T)$ приводят к одной и той же квадратичной форме, поэтому при изучении квадратичных форм достаточно ограничиться только формами, связанными с симметричными матрицами.

Пример 6.8 (Симметричная матрица смежности графа). Рассмотрим неориентированный граф

$$G = \{V, P\},$$

состоящий из множества вершин V и множества P неупорядоченных пар вершин, называемых ребрами:

$$P = \Big\{ \{i_1, j_1\}, \{i_2, j_2\}, \ldots \Big\}.$$

Этот граф можно описать так называемой *матрицей смежности* $A = (a_{ij})$, где

$$a_{ij} = \begin{cases} 1, \text{ если } \{i, j\} \in P, \\ 0 \text{ в противном случае.} \end{cases}$$

Поскольку G — неориентированный граф, вещественная матрица A будет симметричной, т.е. $A^T = A$.

Вещественная симметричная матрица — эрмитова, а значит, нормальная. Комплексная симметричная матрица может не обладать этим свойством.

Пример 6.9 (Симметричность и нормальность).
a) матрица

$$A = \begin{bmatrix} i & 3 & 1-2i \\ 3 & -1 & 3i \\ 1-2i & 3i & 0 \end{bmatrix}$$

является симметричной, но не является эрмитовой, так как

$$A \neq A^* = \begin{bmatrix} -i & 3 & 1+2i \\ 3 & -1 & -3i \\ 1+2i & -3i & 0 \end{bmatrix};$$

b) из двух комплексных симметричных матриц

$$A = \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix}, \quad B = \begin{bmatrix} i & i \\ i & 1 \end{bmatrix}$$

матрица A является нормальной, так как

$$AA^* = \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix} \begin{bmatrix} 1 & -i \\ -i & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} =$$

$$= \begin{bmatrix} 1 & -i \\ -i & 1 \end{bmatrix} \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix} = A^* A,$$

а матрица B не является нормальной, так как

$$BB^* = \begin{bmatrix} i & i \\ i & 1 \end{bmatrix} \begin{bmatrix} -i & -i \\ -i & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1+i \\ 1-i & 2 \end{bmatrix} \neq$$

$$\neq \begin{bmatrix} 2 & 1-i \\ 1+i & 2 \end{bmatrix} = \begin{bmatrix} -i & -i \\ -i & 1 \end{bmatrix} \begin{bmatrix} i & i \\ i & 1 \end{bmatrix} = B^* B.$$

Таким образом, имеется существенное различие между вещественными симметричными и комплексными симметричными матрицами.

Теорема 6.7. *Пусть задана матрица $A \in \mathbb{C}^{n \times n}$. Для того чтобы существовала унитарная матрица $U \in \mathbb{C}^{n \times n}$ и верхняя треугольная матрица $T \in \mathbb{C}^{n \times n}$ такие, что*

$$A = UTU^T, \tag{6.17}$$

необходимо и достаточно, чтобы собственные значения матрицы $A\overline{A}$ *были неотрицательны. При этом условии можно подобрать верхнетреугольную матрицу* T *с неотрицательными диагональными элементами.*

Доказательство**. Необходимость.* Пусть верно (6.17). Тогда

$$A\bar{A} = UTU^T\overline{(UTU^T)} = UTU^T\bar{U}\bar{T}U^* = UT\bar{T}U^*, \tag{6.18}$$

т.е. $A\bar{A}$ унитарно подобна верхнетреугольной матрице $T\bar{T}$ с неотрицательными диагональными элементами (при любой верхнетреугольной матрице T). Следовательно, все собственные значения матрицы $A\bar{A}$ неотрицательные.

Достаточность. Пусть все собственные значения матрицы $A\bar{A}$ неотрицательные и пусть x — ее собственный вектор, т.е.

$$A\overline{A}x = \lambda x, \ \ \lambda \geq 0. \tag{6.19}$$

Возможны две ситуации.

1) Векторы $A\bar{x}$ и x линейно зависимы, т.е. $\exists \mu \in \mathbb{C}, \ \mu \neq 0 :$ $A\bar{x} = \mu x$.

2) Векторы $A\bar{x}$ и x линейно независимы, т.е.

$$y = A\bar{x} + \mu x \neq 0, \ \forall \mu \in \mathbb{C}. \tag{6.20}$$

Покажем, что в ситуации 1) верно $|\mu|^2 = \lambda$:

$$\lambda x = A\bar{A}x = A\overline{(A\bar{x})} = A\,\bar{\mu}\,\bar{x} = \bar{\mu}\,A\,\bar{x} = \bar{\mu}\,\mu\,x = |\mu|^2\,x \Rightarrow \lambda = |\mu|^2\,.$$

Рассмотрим ситуацию 2). Выберем μ такое, что $|\mu|^2 = \lambda$. Тогда

$$A\bar{y} = A\left(\bar{A}x + \bar{\mu}\,\bar{x}\right) = A\bar{A}x + \bar{\mu}\,A\,\bar{x} \xlongequal{(6.19)} \lambda x + \bar{\mu}\,A\,\bar{x} \xlongequal{\lambda=|\mu|^2}$$

$$= \mu\bar{\mu}x + \bar{\mu}\,A\,\bar{x} = \bar{\mu}\,(\mu x + \,A\,\bar{x}) \xlongequal{(6.20)} \bar{\mu}y.$$

Положим в случае 1) $w = x$, $a = \mu$, в случае 2) $w = y$, $a = \bar{\mu}$. Тогда в обоих случаях $\exists w \in \mathbb{C}^n, \ a \in \mathbb{C}, \ |a|^2 = \lambda$ такие, что

$$A\bar{w} = aw. \tag{6.21}$$

Не ограничивая общности w можно считать единичным.

Умножим равенство (6.21) на $e^{-i\Theta}$, где $\Theta \in \mathbb{R}$ — параметр:

$$e^{-i\Theta}A\bar{w} = e^{-i\Theta}aw \ \Leftrightarrow \ A\,\overline{(e^{i\Theta}w)} = \left(e^{-2i\Theta}a\right)\left(e^{i\Theta}w\right).$$

Обозначим $v = e^{i\Theta}w, \ \ \sigma = e^{-2i\Theta}a$. Тогда последнее равенство запишется в виде

$$A\bar{v} = \sigma v.$$

Заметим, что v — единичный вектор для любого $\Theta \in \mathbb{R}$. Выберем Θ так, чтобы $\sigma = e^{-2i\Theta}a \geq 0$. Тогда

$$|\sigma|^2 = \left|e^{-2i\Theta}a\right|^2 = |a|^2 = \lambda.$$

Таким образом, мы показали, что для заданной матрицы A и неотрицательного собственного значения λ матрицы $A\bar{A}$ существует такой единичный вектор v, что

$$A\bar{v} = \sigma v, \ \sigma = \sqrt{\lambda} \geq 0.$$

Дополним этот вектор v до ортонормированного базиса $\{v, v_2, \ldots, v_n\}$ пространства $\mathbb{C}^n$ и составим из векторов этого базиса унитарную матрицу V_1.

Рассмотрим матрицу $V_1^* A\bar{V}_1$. i-й элемент ее первого столбца равен $v_i^* A\bar{v} = \sigma v_i^* v$. Поскольку векторы v и v_i ортонормированны, то справедливо $v_i^* A\bar{v} = \sigma v_i^* v = \sigma\delta_{i1}$, где

$$\delta_{i1} = \begin{cases} 1, \text{ если } i = 1, \\ 0, \text{ если } i \neq 1, \end{cases}$$

— символ Кронекера. Таким образом,

$$V_1^* A\overline{V_1} = \begin{bmatrix} \sigma & \omega^T \\ 0 & A_2 \end{bmatrix}, \ \omega \in \mathbb{C}^{n-1}, \ A_2 \in \mathbb{C}^{(n-1)\times(n-1)}, \ \sigma \geq 0.$$

Тогда

$$\left(V_1^* A\overline{V_1}\right)\overline{\left(V_1^* A\overline{V_1}\right)} = V_1^* A\overline{V_1}V_1^T\bar{A}V_1 \xlongequal{\overline{V_1}V_1^T=E}$$

$$= V_1^* A\bar{A} V_1 = \begin{bmatrix} \sigma^2 & \sigma\omega^T + \omega^T A_2 \\ 0 & A_2\overline{A_2} \end{bmatrix}.$$

Таким образом, множество собственных значений матрицы $A\bar{A}$ состоит из σ^2 и собственных значений матрицы $A_2\overline{A_2}$, следовательно, собственные значения матрицы $A_2\overline{A_2}$ неотрицательны.

Продолжая процесс редукции не более $n-1$ раз (как в доказательстве теоремы Шура), получим

$$V_{n-1}^* \dots V_2^* V_1^* A \overline{V_1 V_2} \dots \overline{V_{n-1}} = \begin{bmatrix} \sigma & * & * \\ 0 & \ddots & * \\ 0 & 0 & \sigma_n \end{bmatrix} = T.$$

Полагая $U = V_1 V_2 \dots V_{n-2}$, имеем

$$U^* A\bar{U} = T \;\Rightarrow\; UTU^T = A. \qquad \square$$

Пример 6.10 (Свойства симметричной матрицы). Покажем, что если матрица A симметрична, то собственные значения матрицы $A\overline{A}$ неотрицательны. Действительно, так как $A = A^T$, то $A\overline{A} = A\overline{A^T} = AA^*$.

Пусть $x \neq 0$ — произвольный собственный вектор матрицы $A\overline{A}$, λ — соответствующее собственное значение, т.е.

$$A\overline{A}x = AA^*x = \lambda x.$$

Тогда

$$x^*\lambda x = x^*AA^*x = (A^*x)^*A^*x,$$

и так как

$$x^*x > 0, \ (A^*x)^*A^*x \geq 0$$

для $x \neq 0$, то

$$\lambda = \frac{(A^*x)^*A^*x}{x^*x} \geq 0.$$

Теорема 6.8 (разложение Такаги). *Матрица $A \in \mathbb{C}^{n\times n}$ симметрична тогда и только тогда, когда существует унитарная матрица $U \in \mathbb{C}^{n\times n}$ и неотрицательная диагональная матрица $\Sigma = \mathrm{diag}(\sigma_1, \dots, \sigma_n)$, такие, что*

$$A = U\Sigma U^T. \tag{6.22}$$

Столбцы матрицы U образуют систему ортонормированных собственных векторов матрицы $A\overline{A}$, а соответствующие диагональные элементы матрицы Σ являются сингулярными числами матрицы A, отвечающими этим собственным векторам.

Доказательство. Во-первых, в примере 6.10 показано, что для любой симметричной матрицы A собственные значения матрицы $A\bar{A}$ неотрицательны. Тогда в силу теоремы 6.7 существует унитарная матрица $U \in \mathbb{C}^{n\times n}$ и верхняя треугольная матрица $T \in \mathbb{C}^{n\times n}$ такие, что

$$A = UTU^T.$$

Но тогда

$$UTU^T = A = A^T = UT^TU^T,$$

следовательно, треугольная матрица T — симметричная, а значит, диагональная. Кроме того, по теореме 6.7 она неотрицательная. Обозначим $\Sigma = T$. Тогда первая часть теоремы доказана.

Докажем вторую часть теоремы. Во-первых, поскольку A симметрична, то $A\bar{A} = A\bar{A}^T = AA^*$ — эрмитова. Рассмотрим разложение матрицы $A\bar{A}$:

$$A\bar{A} \overset{(6.22)}{=} U\Sigma U^T \overline{(U\Sigma U^T)} = U\Sigma U^T \bar{U}\Sigma U^* \overset{U^T\bar{U}=E}{=\!=\!=\!=\!=} U\Sigma^2 U^*.$$

Это представление является унитарной диагонализацией эрмитовой матрицы $A\bar{A}$, и, как доказано в теореме 5.7, столбцы матрицы U образуют систему ортонормированных собственных векторов матрицы $A\bar{A}$.

Так как матрицы $A\bar{A}$ и Σ^2 подобны, то их спектры совпадают. Поэтому и с учетом симметричности матрицы A справедливы соотношения

$$\sigma(\Sigma^2) = \sigma(A\bar{A}) = \sigma(A\bar{A^T}) = \sigma(AA^*),$$

т.е. диагональные элементы матрицы Σ являются сингулярными числами матрицы A. □

Замечание 6.1. Условие представления матрицы A в виде (6.22) является также достаточным для симметричности матрицы A, поскольку любая матрица $A = U\Sigma U^T$ с диагональной матрицей Σ, очевидно, симметрична.

6.4. Знакоопределенные матрицы

Определение 6.3. Эрмитова $n \times n$-матрица называется *положительно определенной*, если

$$x^*Ax > 0, \ \forall x \in \mathbb{C}^n, \ x \neq 0. \tag{6.23}$$

Если в (6.23) знак $>$ заменить на $\geq$ $(<, \leq)$, получим определение *положительно полуопределенной (отрицательно определенной, отрицательно полуопределенной)* матрицы.

Если эрмитова матрица не принадлежит ни к одному из указанных выше классов, то ее называют *незнакоопределенной*.

Теорема 6.9 (Критерий Сильвестра). *[1, стр. 101], [7, стр. 285-289]. Для того чтобы эрмитова матрица была положительно определенной, необходимо и достаточно, чтобы все главные миноры этой матрицы были положительны.*

Теорема 6.10 (Критерий Якоби). *[5, стр. 88]. Для того чтобы эрмитова матрица была положительно определенной, необходимо и достаточно, чтобы все коэффициенты характеристического полинома были отличны от нуля и имели чередующиеся знаки.*

Свойства знакоопределенных матриц

1^0. Если $A \in \mathbb{C}^{n\times n}$ положительно определена, то это же верно для матриц $\overline{A}$, A^T, A^*.

2^0. Матрица $A \in \mathbb{C}^{n\times n}$ положительно определена тогда и только тогда, когда матрица A^{-1} существует и положительно определена.

Определение 6.4. *Главная* $k \times k$*-подматрица в* $A \in \mathbb{C}^{n\times n}$ — это подматрица, расположенная на пересечении k строк и k столбцов с одинаковыми множествами номеров. Например, для матрицы

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

главными подматрицами будут

$$A_{13} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \ A_{12} = \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix}.$$

3^0. Всякая главная подматрица положительно определенной матрицы сама положительно определена.

4^0. Пусть $A, B \in \mathbb{C}^{n\times n}$ — положительно полуопределенные матрицы. Тогда для $\forall \alpha, \beta > 0$ матрица $\alpha A + \beta B$ положительно полуопределена.

5^0. Каждое собственное значение положительно определенной (полуопределенной) матрицы положительно (неотрицательно).

6^0. След, определитель и все главные миноры положительно определенной матрицы положительны.

7^0. Для любой матрицы $A \in \mathbb{C}^{n\times n}$ матрицы A^*A и AA^* положительно полуопределенные, собственные значения матриц A^*A и AA^* неотрицательны.

Доказательство. 1^0. Пусть A положительно определена, т.е. для нее справедливо (6.23). Докажем положительную определенность матрицы $\overline{A}$, т.е. покажем, что (6.23) верно также и с матрицей $\overline{A}$ вместо A:

$$x^*\overline{A}x = \overline{\overline{x}^*A\overline{x}} = \left[y \stackrel{\Delta}{=} \overline{x}\right] = \overline{y^*Ay}. \tag{6.24}$$

Так как (6.23) верно для любого $x \in \mathbb{C}^n$, $x \neq 0$, то оно верно и для $y = \overline{x} \neq 0$. Таким образом, $y^*Ay > 0$, следовательноб и $\overline{y^*Ay} > 0$, и согласно (6.24) имеем $x^*\overline{A}x > 0$, что и доказывает положительную определенность $\overline{A}$.

Для матриц A^T, A^* доказательство проводится аналогично.

2^0. Во-первых, положительно определенная матрица A обязательно невырожденная, т.к. если матрица вырожденная, то существует вектор $x \neq 0$, для которого $Ax = 0$, и следовательно, $x^*Ax = 0$, что противоречит определению положительно определенной матрицы.

Докажем теперь, что A^{-1} положительно определенная. Для этого преобразуем (6.23) для матрицы A^{-1} с учетом очевидного равенства:

$$E = (A^*)^{-1}A^* = (A^{-1})^*A^* \tag{6.25}$$

и свойства 1^0 для A^* :

$$x^*A^{-1}x = x^*(A^{-1})^*A^*A^{-1}x = (A^{-1}x)^*A^*A^{-1}x =$$
$$= \left[y \stackrel{\Delta}{=} A^{-1}x\right] = y^*A^*y > 0.$$

3^0. Для собственного подмножества S, состоящего из m элементов множества $\{1, 2, ...n\}$, рассмотрим главную подматрицу $A(S)$ со строками и столбцами из S, произвольный вектор $x \in \mathbb{C}^m$, и вектор $y = [y_1, ..., y_n]^T \in \mathbb{C}^n$: $y_i = 0$, $\forall i \in S$, $y_i = x_i$, $i \notin S$:

$$x^*A(S)x = y^*Ay > 0, \ \forall x \in \mathbb{C}^m, \ x \neq 0.$$

4^0. Легко следует из справедливости неравенства (6.23) для матриц A и B.

5^0. Пусть $\lambda \in \sigma(A)$ и

$$Ax = \lambda x, \ x \neq 0. \tag{6.26}$$

Домножим (6.26) слева на x^*:

$$x^*Ax = x^*\lambda x = \lambda x^*x \Rightarrow \lambda = \frac{x^*Ax}{x^*x} \stackrel{(6.23),(6.26)}{>} 0.$$

6^0. Следует из 5^0 и теоремы 5.6.

7^0. Положительная полуопределенность матрицы A^*A легко следует из очевидных соотношений

$$x^*A^*Ax = (Ax)^*Ax \geq 0, \ \forall\, x \in \mathbb{C}^n$$

и определения 6.3.

Утверждение для собственных значений доказывается аналогично 5^0. $\square$

Пример 6.11 (Положительно определенная матрица). Покажем, что матрица

$$A = \begin{bmatrix} 5 & -1 & 3 \\ -1 & 2 & -2 \\ 3 & -2 & 3 \end{bmatrix}$$

положительно определена.

Матрица A эрмитова. Согласно критерию Сильвестра вычислим ее главные миноры:

$$\Delta_1 = 5 > 0,$$
$$\Delta_2 = \begin{vmatrix} 5 & -1 \\ -1 & 2 \end{vmatrix} = 8 > 0,$$
$$\Delta_3 = \begin{vmatrix} 5 & -1 & 3 \\ -1 & 2 & -2 \\ 3 & -2 & 3 \end{vmatrix} = 17 > 0,$$

что подтверждает положительную определенность матрицы A.

Запишем характеристический полином матрицы A:

$$p_A(\lambda) = \lambda^3 - 10\lambda^2 + 31\lambda - 30.$$

Согласно критерию Якоби матрица A положительно определена.

Пусть $A \in \mathbb{C}^{m\times n}$, $\operatorname{rank} A = r$. Тогда $A^*A \in \mathbb{C}^{n\times n}$ — эрмитова положительно полуопределенная матрица, все ее собственные значения $\lambda_i(A^*A)$ неотрицательны.

Сингулярные числа $\sigma_i(A) = \sqrt{\lambda_i(A^*A)}$ матрицы A упорядочим по невозрастанию:

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > \sigma_{r+1} = \cdots = \sigma_n = 0. \tag{6.27}$$

Определение 6.5. *Первым сингулярным базисом* матрицы $A \in \mathbb{C}^{m\times n}$ называется ортонормированный базис V пространства $\mathbb{C}^n$ из собственных векторов v_i матрицы A^*A, упорядоченных таким образом, что соответствующие собственные значения не возрастают:

$$A^*Av_i = \begin{cases} \sigma_i^2 v_i, & 1 \leq i \leq r, \\ 0, & r+1 \leq i \leq n. \end{cases} \tag{6.28}$$

Определение 6.6. *Вторым сингулярным базисом* матрицы $A \in \mathbb{C}^{m\times n}$ называется ортонормированный базис $U = \{u_1, \ldots, u_m\}$ пространства $\mathbb{C}^m$, первые r векторов которого имеют вид

$$u_i = \frac{Av_i}{\sigma_i},\ i = \overline{1,r}, \tag{6.29}$$

где $v_1, \ldots, v_r$ — первые r векторов первого сингулярного базиса, $\sigma_1, \ldots, \sigma_r$ — ненулевые сингулярные числа матрицы A.

Для $i, j = \overline{1,r}$:

$$\left\langle \frac{Av_i}{\sigma_i}, \frac{Av_j}{\sigma_j} \right\rangle = \frac{\langle A^*Av_i, v_j\rangle}{\sigma_i\sigma_j} = \begin{cases} 0, & i \neq j, \\ 1, & i = j. \end{cases} \tag{6.30}$$

Это означает, что векторы (6.29) второго сингулярного базиса ортонормированы.

Из определения следует, что сингулярные базисы определяются неоднозначно.

Определение 6.7. *Диагональной прямоугольной матрицей* $A \in \mathbb{C}^{m\times n}$ называется матрица, у которой $a_{ij} = 0$ при $j \neq i$.

Пример 6.12 (Диагональные прямоугольные матрицы). Прямоугольные матрицы

$$A = \left[\begin{array}{cc|ccc} 1 & 0 & 0 & 0 & 0 \\ 0 & 9-3i & 0 & 0 & 0 \end{array}\right] \quad \text{и} \quad B = \left[\begin{array}{ccc} 1 & 0 & 0 \\ 0 & -5i & 0 \\ 0 & 0 & 6i \\ \hline 0 & 0 & 0 \end{array}\right]$$

являются диагональными.

Теорема 6.11 (Сингулярное разложение). *Любая матрица $A \in \mathbb{C}^{m\times n}$ ранга r может быть представлена в виде*

$$A = U\Sigma V^*, \tag{6.31}$$

где $U \in \mathbb{C}^{m\times m}$, $V \in \mathbb{C}^{n\times n}$ — унитарные матрицы, $\Sigma = [\sigma_{ij}] \in \mathbb{C}^{m\times n}$ — диагональная прямоугольная матрица с невозрастающими неотрицательными элементами по диагонали. При этом

— диагональные элементы σ_{ii} являются сингулярными числами матрицы A;

— столбцы матрицы U образуют ортонормированный базис из собственных векторов матрицы AA^*;

— столбцы матрицы V образуют ортонормированный базис из собственных векторов матрицы A^*A.

Доказательство. Согласно свойству 7^0 знакоопределенных матриц эрмитова матрица A^*A является положительно полуопределенной. Пусть $\sigma_1^2, \ldots, \sigma_n^2$ — собственные значения матрицы A^*A, причем σ_i упорядочены как в (6.27). Пусть $v_1, \ldots, v_n$ — собственные векторы матрицы $A^*A \in \mathbb{C}^{n\times n}$, соответствующие $\sigma_1^2, \ldots, \sigma_n^2$. Из спектральной теоремы для нормальных матриц следует, что $v_1, \ldots, v_n$ можно выбрать ортонормированными.

По системе векторов $v_1, \ldots, v_r$ построим второй сингулярный базис $u_1, \ldots, u_m$ (если $r < m$, то систему ортонормированных векторов $u_1, \ldots, u_r$ (6.30) добавим до ортонормированного базиса в m-мерном пространстве).

Составим из векторов $v_1, \ldots, v_n$ и $u_1, \ldots, u_m$ две унитарные матрицы: $V = \{v_1, \ldots, vx_n\} \in \mathbb{C}^{n\times n}$, $U = \{u_1, \ldots, u_m\} \in \mathbb{C}^{m\times m}$. Составим также диагональную матрицу $\Sigma = \operatorname{diag}\{\sigma_1, \ldots, \sigma_n\}$. Тогда согласно (6.30), (6.29)

$$A = U\Sigma V^*. \qquad \square$$

Упражнение 6.4. Доказать, что столбцы матрицы V из сингулярного разложения (6.31) образуют первый сингулярный базис, а столбцы матрицы U — второй сингулярный базис матрицы A.

Пример 6.13 (Сингулярное разложение). Построим сингулярное разложение матрицы

$$A = \begin{bmatrix} 1 & -1 & 2 \\ -1 & 1 & -2 \\ 2 & -2 & 4 \end{bmatrix}.$$

Найдем сингулярные числа матрицы A. Для этого найдем матрицу A^*A :

$$A^*A = \begin{bmatrix} 1 & -1 & 2 \\ -1 & 1 & -2 \\ 2 & -2 & 4 \end{bmatrix} \begin{bmatrix} 1 & -1 & 2 \\ -1 & 1 & -2 \\ 2 & -2 & 4 \end{bmatrix} = \begin{bmatrix} 6 & -6 & 12 \\ -6 & 6 & -12 \\ 12 & -12 & 24 \end{bmatrix}.$$

Вычислим

$$\det(A^*A - \lambda E_3) = \lambda^2(\lambda - 36).$$

Значит, собственные значения матрицы A^*A : $\lambda_1(A^*A) = 36$, $\lambda_2(A^*A) = \lambda_3(A^*A) = 0$, а сингулярные числа матрицы A : $\sigma_1 = 6$, $\sigma_2 = \sigma_3 = 0$. Тогда матрица Σ из сингулярного разложения (6.31) имеет вид:

$$\Sigma = \begin{bmatrix} 6 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Найдем первый сингулярный базис. Он состоит из собственных векторов матрицы A^*A.

Для $\lambda_1 = 36$ имеем:

$$A^*A - 36E_3 = \begin{bmatrix} -30 & -6 & 12 \\ -6 & -30 & -12 \\ 12 & -12 & -12 \end{bmatrix} \sim \begin{bmatrix} -36 & -36 & 0 \\ -6 & -30 & -12 \\ 18 & 18 & 0 \end{bmatrix} \sim$$
$$\sim \begin{bmatrix} 1 & 5 & 2 \\ 1 & 1 & 0 \end{bmatrix} \sim \begin{bmatrix} 0 & 2 & 1 \\ 1 & 1 & 0 \end{bmatrix}.$$

В качестве нормированного собственного вектора можно взять

$$v^{(1)} = \frac{1}{\sqrt{6}} \begin{bmatrix} -1 \\ 1 \\ -2 \end{bmatrix}.$$

Найдем собственные векторы, соответствующие собственному значению $\lambda = 0$. Имеем:

$$A^*A - 0\cdot E_3 = \begin{bmatrix} -6 & -6 & 12 \\ -6 & -6 & -12 \\ 12 & -12 & 24 \end{bmatrix} \sim \begin{bmatrix} -1 & 1 & -2 \end{bmatrix}.$$

Таким образом, пространство собственных векторов, соответствующих двухкратному собственному значению $\lambda = 0$, имеет вид:

$$v = \begin{bmatrix} \alpha - 2\beta \\ \alpha \\ \beta \end{bmatrix}.$$

Возьмем нормированные линейно независимые векторы

$$v^{(2)} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad v^{(3)} = \frac{1}{\sqrt{5}} \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}.$$

Заметим, что полученная система линейно независимых собственных векторов эрмитовой матрицы A^*A, состоящая из $v^{(1)}$, $v^{(2)}$, $v^{(3)}$ ортонормирована.

Таким образом, первый сингулярный базис матрицы A имеет вид:

$$v^{(1)} = \frac{1}{\sqrt{6}} \begin{bmatrix} -1 \\ 1 \\ -2 \end{bmatrix}, \quad v^{(2)} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad v^{(3)} = \frac{1}{\sqrt{5}} \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}.$$

Из векторов первого сингулярного базиса формируем столбцы матрицы V сингулярного разложения (6.31):

$$V = \begin{bmatrix} \frac{-1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{-2}{\sqrt{5}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{-2}{\sqrt{6}} & 0 & \frac{1}{\sqrt{5}} \end{bmatrix}.$$

Найдем второй сингулярный базис. Поскольку $\operatorname{rank} A = 1$, то по формуле (6.29) вычисляем только один вектор второго сингулярного базиса:

$$u^{(1)} = \frac{Ax^{(1)}}{\sigma_1} = \frac{1}{6} \begin{bmatrix} 1 & -1 & 2 \\ -1 & 1 & -2 \\ 2 & -2 & 4 \end{bmatrix} \frac{1}{\sqrt{6}} \begin{bmatrix} -1 \\ 1 \\ -2 \end{bmatrix} = \frac{1}{\sqrt{6}} \begin{bmatrix} -1 \\ 1 \\ -2 \end{bmatrix}.$$

Дополняем $u^{(1)}$ до ортонормированного базиса произвольными векторами (в общем случае можно дополнить до базиса любыми линейно независимыми векторами, а затем применить к полученной системе векторов процедуру ортогонализации Грама—Шмидта (см. стр. 135)):

$$u^{(2)} = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix}, \quad u^{(3)} = \frac{1}{\sqrt{10}} \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}.$$

Таким образом, второй сингулярный базис матрицы A имеет вид:

$$u^{(1)} = \frac{1}{\sqrt{6}} \begin{bmatrix} -1 \\ 1 \\ -2 \end{bmatrix}, \quad u^{(2)} = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix}, \quad u^{(3)} = \frac{1}{\sqrt{10}} \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}.$$

Из векторов второго сингулярного базиса формируем столбцы матрицы U сингулярного разложения (6.31):

$$U = \begin{bmatrix} \frac{-1}{\sqrt{6}} & \frac{2}{\sqrt{5}} & 0 \\ \frac{1}{\sqrt{6}} & 0 & \frac{2}{\sqrt{10}} \\ \frac{-2}{\sqrt{6}} & \frac{-1}{\sqrt{5}} & \frac{1}{\sqrt{10}} \end{bmatrix}.$$

Окончательно имеем сингулярное разложение матрицы A:

$$U\Sigma V^* = \begin{bmatrix} \frac{-1}{\sqrt{6}} & \frac{2}{\sqrt{5}} & 0 \\ \frac{1}{\sqrt{6}} & 0 & \frac{2}{\sqrt{10}} \\ \frac{-2}{\sqrt{6}} & \frac{-1}{\sqrt{5}} & \frac{1}{\sqrt{10}} \end{bmatrix} \begin{bmatrix} 6 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{-1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{-2}{\sqrt{5}} & 0 & \frac{1}{\sqrt{5}} \end{bmatrix} =$$

$$= \begin{bmatrix} 1 & -1 & 2 \\ -1 & 1 & -2 \\ 2 & -2 & 4 \end{bmatrix} = A. \tag{6.32}$$

Теорема 6.12 (Полярное разложение). *Если $A \in \mathbb{C}^{n\times n}$, то существуют положительно полуопределенные эрмитовы матрицы H и K, однозначно определяемые матрицей A, и унитарные матрицы W и Y из $\mathbb{C}^{n\times n}$, такие что*

$$A = WH = KY. \tag{6.33}$$

Более того, $H = (A^*A)^{\frac{1}{2}}$, $K = (AA^*)^{\frac{1}{2}}$.

Доказательство. Положим

$$W = UV^*, \quad H = V\Sigma V^*, \tag{6.34}$$

где U, V — матрицы из сингулярного разложения (6.31). Несложно проверить, что W — унитарная матрица, а H — положительно полуопределенная эрмитова матрица. Тогда имеем:

$$WH = UV^*V\Sigma V^* = U\Sigma V^* = A.$$

Далее, так как $A^*A = HW^*WH = H^2$, то верно и $H = (A^*A)^{\frac{1}{2}}$.

Применяя полученный результат к A^*, получим $A = KY$, $K = (AA^*)^{\frac{1}{2}}$. $\square$

Упражнение 6.5. Доказать, что матрицы W и Y из полярного разложения (6.33) определяются единственным образом тогда и только тогда, когда A невырожденная, а также, что если $A \in \mathbb{R}^{n\times n}$, то все матрицы могут быть выбраны действительными.

Пример 6.14 (Сингулярное и полярное разложение). Построим сингулярное и полярное разложение для матрицы

$$A = \begin{bmatrix} -2 & -4 \\ 1 & 2 \end{bmatrix}. \tag{6.35}$$

Найдем сингулярные числа матрицы A. Для этого найдем матрицу A^*A :

$$A^*A = \begin{bmatrix} -2 & 1 \\ -4 & 2 \end{bmatrix} \begin{bmatrix} -2 & -4 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 5 & 10 \\ 10 & 20 \end{bmatrix}.$$

Вычислим

$$\det(A^*A - \lambda E_2) = \lambda(\lambda - 25).$$

Значит, собственные значения матрицы A^*A : $\lambda_1(A^*A) = 25$, $\lambda_2(A^*A) = 0$, а сингулярные числа матрицы A : $\sigma_1 = 5$, $\sigma_2 = 0$. Значит, матрица Σ из сингулярного разложения (6.31) для матрицы (6.35) имеет вид:

$$\Sigma = \begin{bmatrix} 5 & 0 \\ 0 & 0 \end{bmatrix}.$$

Найдем первый сингулярный базис. Он состоит из собственных векторов матрицы A^*A.

Для $\lambda_1 = 25$ имеем:

$$A^*A - 25E_3 = \begin{bmatrix} -20 & 10 \\ 10 & -5 \end{bmatrix} \sim \begin{bmatrix} -2 & 1 \end{bmatrix}.$$

В качестве нормированного собственного вектора можно взять

$$v^{(1)} = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

Найдем собственный вектор, соответствующий собственному значению $\lambda = 0$. Имеем:

$$A^*A - 0 \cdot E_2 = \begin{bmatrix} 5 & 10 \\ 10 & 20 \end{bmatrix} \sim \begin{bmatrix} -2 & 1 \end{bmatrix}.$$

В качестве второго нормированного собственного вектора возьмем

$$v^{(2)} = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

Таким образом, первый сингулярный базис матрицы A имеет вид:

$$v^{(1)} = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad v^{(2)} = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

Из векторов первого сингулярного базиса формируем столбцы матрицы V сингулярного разложения (6.31):

$$V = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix}.$$

Найдем второй сингулярный базис. Поскольку $\operatorname{rank} A = 1$, то по формуле (6.29) вычисляем только один вектор второго сингулярного базиса:

$$u^{(1)} = \frac{1}{\sigma_1} A x^{(1)} = \frac{1}{5} \begin{bmatrix} -2 & -4 \\ 1 & 2 \end{bmatrix} \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \frac{1}{\sqrt{5}} \begin{bmatrix} -2 \\ 1 \end{bmatrix}.$$

Дополняем $u^{(1)}$ до ортонормированного базиса:

$$u^{(2)} = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

Таким образом, второй сингулярный базис матрицы A имеет вид:

$$u^{(1)} = \frac{1}{\sqrt{5}} \begin{bmatrix} -2 \\ 1 \end{bmatrix}, \quad u^{(2)} = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

Из векторов второго сингулярного базиса формируем столбцы матрицы U сингулярного разложения (6.31):

$$U = \frac{-1}{\sqrt{6}} \begin{bmatrix} -2 & 1 \\ 1 & 2 \end{bmatrix}.$$

Окончательно имеем сингулярное разложение матрицы A:

$$U\Sigma V^* = \frac{-1}{\sqrt{6}} \begin{bmatrix} -2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 5 & 0 \\ 0 & 0 \end{bmatrix} \frac{1}{\sqrt{5}} \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix} = \begin{bmatrix} -2 & -4 \\ 1 & 2 \end{bmatrix} = A. \tag{6.36}$$

Построим полярное разложение (6.33), используя (6.34):

$$W = UV^* = \frac{-1}{\sqrt{6}} \begin{bmatrix} -2 & 1 \\ 1 & 2 \end{bmatrix} \frac{1}{\sqrt{5}} \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix} = \frac{1}{5} \begin{bmatrix} -4 & -3 \\ -3 & 4 \end{bmatrix},$$

$$H = W^* A = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix} \frac{1}{\sqrt{5}} \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}.$$

В итоге получаем полярное разложение:

$$WH = \frac{1}{5} \begin{bmatrix} -4 & -3 \\ -3 & 4 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \frac{1}{\sqrt{5}} \begin{bmatrix} -10 & -20 \\ 5 & 10 \end{bmatrix} = \begin{bmatrix} -2 & -4 \\ 1 & 2 \end{bmatrix} = A. \tag{6.37}$$

Задачи и упражнения

1. Доказать, что если $A, B \in \mathbb{C}^{n \times n}$ — эрмитовы, то AB эрмитова тогда и только тогда, когда A и B коммутируют (перестановочны).
2. Записать эрмитово разложение матрицы
$$A = \begin{bmatrix} i & -2 \\ 3i & 1 \end{bmatrix}.$$
3. Доказать, что определитель эрмитовой матрицы есть вещественное число.
4. Доказать, что все собственные значения косоэрмитовой матрицы — чисто мнимые.
5. Доказать, что столбцы матрицы U из (6.4) являются ортонормированными собственными векторами матрицы A.
6. Записать спектральное разложение матрицы
$$A = \begin{bmatrix} 1 & 0 & -4 \\ 0 & 5 & 4 \\ -4 & 4 & 3 \end{bmatrix}.$$

7. Доказать, что если U унитарна и эрмитова, то $U^2 = E$ (т.е. матрица U *инволютивна*).

8. Пусть заданы эрмитова матрица $A \in \mathbb{C}^{n\times n}$ и ненулевой вектор $x \in \mathbb{C}^n$, и пусть $\alpha = \dfrac{x^*Ax}{x^*x}$. Доказать, что тогда в каждом из полуинтервалов $(-\infty, \alpha]$ и $[\alpha, \infty)$ найдется по крайней мере одно собственное значение матрицы A.

9. Для собственного вектора x с собственным значением λ эрмитовой матрицы A показать, что $\dfrac{x^*Ax}{x^*x} = \lambda$.

10. Дана эрмитова матрица
$$A = \begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix}.$$
Найти экстремальные значения отношения Рэлея—Ритца $g(x) = \dfrac{x^TAx}{x^Tx}$, $x \neq 0,\ x \in \mathbb{R}^2$.

11*. Пусть задана матрица $A \in \mathbb{C}^{n\times n}$ с собственными значениями $\{\lambda_i\}$. Доказать, что верны оценки
$$|\lambda_i| \le \max_{x\neq 0}\left|\frac{x^*Ax}{x^*x}\right|, \quad \min_{x\neq 0}\left|\frac{x^*Ax}{x^*x}\right| \le |\lambda_i|, \quad i = \overline{1,n},$$
даже если матрица не является эрмитовой.

На примере $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ удостовериться, что каждая из этих оценок может быть грубой.

12. На примере матриц
$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \quad \text{и} \quad B = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$
показать, что неравенства из теоремы Вейля могут потерять силу, если матрицы A и B не являются эрмитовыми.

13. Задана последовательность комплексных чисел $\{a_0, a_1, \ldots\}$ и натуральное число n. Положим $A_{2n} = [a_{ij}] \equiv [a_{i+j}] \in \mathbb{C}^{2n\times 2n}$ (*ганкелева матрица* — матрица, элементы a_{ij} которой суть функции только от суммы индексов $i + j$). Убедиться, что A_{2n} — комплексная симметричная матрица. Какое количество значений надо определить, чтобы задать ганкелеву матрицу?

14. Пусть $x \in \mathbb{R}^n$ и $f(x)$ — комплекснозначная дважды непрерывно дифференцируемая функция, определенная в некоторой области $D \subset \mathbb{R}^n$. Рассмотрим линейный оператор L второго порядка с частными производными, определенный выражением
$$Lf(x) \equiv \sum_{i,j=1}^{n} a_{ij}(x)\frac{\partial^2 f(x)}{\partial x_i \partial x_j}, \tag{6.38}$$
и соответствующую ему матрицу коэффициентов $A = (a_{ij}(x))$. Показать, что при изучении вещественных или комплексных линейных дифференциальных операторов с частными производными вида (6.38) без ограничения общности можно рассматривать лишь случай симметричной матрицы коэффициентов.

15. Пусть $A \in \mathbb{C}^{n\times n}$ симметрична, $A = B + iC$, $B, C \in \mathbb{R}^{n\times n}$. Показать, что матрица A нормальна тогда и только тогда, когда выполнено любое из утверждений

a) B и C — перестановочны,

b) $A\bar{A}$ — вещественная,

c) A и $\overline{A}$ — перестановочны.

16. Привести пример симметричной матрицы, не являющейся нормальной.

17. Выяснить, какие из нижеприведенных матриц A могут быть триангуляризованы преобразованием вида $A \to UTU^T$ с унитарной матрицей $U \in \mathbb{C}^{n\times n}$:

a) $\begin{bmatrix} 1 & i \\ -i & 1 \end{bmatrix}$, b) $\begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix}$, c) $\begin{bmatrix} 1 & i \\ 1 & -1 \end{bmatrix}$, d) $\begin{bmatrix} 1 & 0 \\ 1 & -i \end{bmatrix}$,
e) $\begin{bmatrix} 1 & 0 \\ 0 & i \end{bmatrix}$.

18. Пусть $A = U\Sigma U^T \in \mathbb{C}^{n\times n}$, где матрицы U и Σ из разложения Такаги. Проверить прямыми вычислениями, что числа σ_i^2

(квадраты диагональных элементов матрицы Σ) являются собственными значениями матриц $A\bar{A}$ и $\bar{A}A$ и что эти матрицы эрмитовы. Показать, что столбцы u_i матрицы U и числа σ_i удовлетворяют условиям $A\bar{u}_i = \sigma_i u_i$, $i = \overline{1,n}$.

19. Пусть матрица $A \in \mathbb{C}^{n\times n}$ симметрична и задано разложение $A = U\Sigma U^T$ с унитарной матрицей U и диагональной матрицей $\Sigma = \operatorname{diag}[\sigma_1, \sigma_2, \ldots, \sigma_n]$, где все числа σ_i больше или равны 0. Показать, что rank A равен числу ненулевых значений σ_i.

20. Пусть для матрицы $A \in \mathbb{C}^{n\times n}$ существует такая невырожденная матрица $S \in \mathbb{C}^{n\times n}$, что $A = S\Lambda\bar{S}^{-1}$, где $\Lambda = \operatorname{diag}[\lambda_1, \lambda_2, \ldots, \lambda_n]$. Убедиться, что матрица $A\bar{A}$ диагонализуема, все ее собственные значения неотрицательны и $\operatorname{rank} A = \operatorname{rank} A\bar{A}$. Какое отношение это имеет к разложению Такаги? Проверить, что ни одна из матриц
$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$
не может быть записана в указанной форме.

21. Показать, что собственные векторы $x, y \in \mathbb{C}^n$ комплексной симметричной матрицы $A \in \mathbb{C}^{n\times n}$, отвечающие различным собственным значениям, удовлетворяют соотношению $x^T y = 0$. Означает ли это ортогональность векторов x и y? Привести пример.

22. Вещественная симметричная матрица эрмитова, и, следовательно, диагонализуема. Показать на примере матрицы
$$A = \begin{bmatrix} 1 & i \\ i & -1 \end{bmatrix},$$
что комплексная симметричная матрица может не быть диагонализуемой.

23. Что означают свойства положительной определенности и положительной полуопределенности при $n = 1$?

24. Показать, что собственные значения и след отрицательно определенной матрицы порядка n отрицательны, а определитель отрицателен для нечетных n и положителен для четных.

25. Показать, что собственные значения, след, определитель и главные миноры положительно полуопределенной матрицы суть неотрицательные числа.

26. Доказать, что положительность (неотрицательность) всех собственных значений эрмитовой матрицы достаточна для того, чтобы она была положительно (неотрицательно) определенной.

27. Доказать, что если $A \in \mathbb{C}^{n\times n}$ положительно полуопределена, то положительно полуопределены и все степени A^k, $k = 1, 2, \ldots$.

28. Показать, что ведущие главные миноры симметричной матрицы $\left[\begin{smallmatrix} 0 & 0 \\ 0 & -1 \end{smallmatrix}\right]$ неотрицательны, но матрица не является положительно полуопределенной.

29. Выяснить, является ли матрица положительно определенной:
$$A=\begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix}, B=\begin{bmatrix} 1 & 3 \\ 3 & 5 \end{bmatrix}, C=\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix},$$
$$D=\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}, F=\begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}, G=\begin{bmatrix} 2 & 0 & 1 \\ 0 & 2 & 1 \\ 1 & 1 & 3 \end{bmatrix}.$$

30. Доказать, что если $A = U\Sigma V^*$ — сингулярное разложение матрицы A, то $A^+ = V\Sigma^+ U^*$, где Σ^+ получена из Σ транспонированием и заменой положительных диагональных элементов обратными величинами.

31. Доказать свойства (a)—(n) из упражнения 6 стр. 54 , используя сингулярное разложение.

32. Показать, что если $A \in \mathbb{C}^{n\times n}$ — симметричная и невырожденная, то обратная к ней A^{-1} также симметрична.

33. Пусть $A \in \mathbb{C}^{n\times n}$. Доказать, что матрица A симметрична тогда и только тогда, когда существует матрица $S \in \mathbb{C}^{n\times n}$, такая, что $A = SS^T$.

34*. На примере матрицы $A = \left[\begin{smallmatrix} 1 & 2 \\ 0 & 1 \end{smallmatrix}\right]$, показать, что предположение теоремы Рэлея—Ритца о том, что матрица A — эрмитова, является существенным.

35*. Показать, что матрица U_w из примера 5.5 одновременно унитарна и эрмитова.

36. Доказать, что классы унитарных, эрмитовых, косоэрмитовых, нормальных матриц замкнуты относительно преобразования унитарного подобия.

37. Почему не является полярным разложением равенство

$$\begin{bmatrix} 5 & 0 \\ 4 & 5 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 4 & 5 \\ 5 & 0 \end{bmatrix}?$$

38. Построить сингулярное и полярное разложение матриц

$$A = \begin{bmatrix} 11 & 10 \\ -2 & 5 \end{bmatrix},\ B = \begin{bmatrix} 1+i & 1-i \\ 1+i & -1+i \end{bmatrix},\ C = \begin{bmatrix} \sqrt{3} & -2 \\ 0 & \sqrt{3} \end{bmatrix},$$
$$D = \begin{bmatrix} 7 & 7 \\ 1 & 1 \end{bmatrix}, E= \begin{bmatrix} 0 & 0 \\ 4 & -3 \end{bmatrix},\ F = \begin{bmatrix} \sqrt{2} & 1 \\ 0 & \sqrt{2} \end{bmatrix},\ G = \begin{bmatrix} 5 & 0 \\ 4 & 3 \end{bmatrix}.$$

Вопросы для повторения

1. Пусть $A \in \mathbb{C}^{n\times n}$ — произвольная матрица. Какие из следующих матриц являются эрмитовыми, косоэрмитовыми: a)$A + A^*$, b) $A - A^*$, c) AA^*, d) A^*A, e) $S^*AS,\ S \in \mathbb{C}^{n\times n}$?

2. Пусть $A, B \in \mathbb{C}^{n\times n}$ — эрмитовы (косоэрмитовы). Какие из следующих матриц являются эрмитовыми (косоэрмитовыми): a) A^k; в) A^{-1}, $(\det A \neq 0)$; c) iA; d) $\alpha A + \beta B$, $\alpha, \beta \in \mathbb{R}$; e) $A + iB$?

3. Что можно сказать о диагональных и внедиагональных элементах эрмитовой матрицы? Косоэрмитовой матрицы?

4. Для какой косоэрмитовой матрицы определитель есть вещественное число? Определитель имеет нулевую вещественную часть?

5. Пусть $A \in \mathbb{C}^{n\times n}$ — эрмитова (косоэрмитова) матрица. Что можно сказать о значении x^*Ax для любого $x \in \mathbb{C}^n$? О собственных значениях матрицы A?

6. Дайте вариационные описания собственных значений $\lambda_1, \ldots, \lambda_n$ эрмитовой матрицы.

7. Какая матрица называется симметричной?

8. При каких условиях симметричная матрица является эрмитовой?

9. Когда для произвольной матрицы $A \in \mathbb{C}^{n\times n}$ существует унитарная матрица $U \in \mathbb{C}^{n\times n}$ и верхняя треугольная матрица $T \in \mathbb{C}^{n\times n}$, такие, что $A = UTU^T$?

10. Что можно сказать о собственных значениях матрицы $A\bar{A}$, если A симметрична.

11. Сформулируйте свойства знакоопределенных матриц.

12. Дайте определения первого и второго сингулярного базисов.

13. Какая связь между сингулярным и полярным разложениями?

Необходимо усвоить

Определения и свойства матриц

- косоэрмитовая
- незнакоопределенная
- нормальная
- отрицательно определенная
- отрицательно полуопределенная
- положительно определенная
- положительно полуопределенная
- симметричная
- эрмитова

Формулировку и трактовку теорем

- об эрмитовом разложении
- критерии эрмитовости матрицы
- спектральное разложение эрмитовых матриц
- Теорема Рэлея-Ритца
- Теорема Вейля
- разложение Такаги
- о сингулярном разложении
- о полярном разложении

Основные умения и навыки

- определять принадлежность заданной матрицы определенному классу
- распознавать по матрице возможность построения для нее того или иного разложения и строить (если возможно) это разложение (спектральное, сингулярное, полярное).

Часть III

ПРИКЛАДНЫЕ АСПЕКТЫ ЛИНЕЙНОЙ АЛГЕБРЫ

Глава 7

Нормы векторов и матриц

Во многих приложениях линейной алгебры надо уметь «измерять» векторы и матрицы (например, при решении задачи о влиянии ошибок измерения или округления на решения линейных систем). С этой целью вводятся нормы векторов и матриц. Нормы играют важную роль при анализе и оценивании численных методов, часто привлекаются при выводе оценок для важных величин, связанных с матрицей, например собственных значений, существенно используются при анализе устойчивости алгоритмов вычисления линейной алгебры, при изучении матричных степенных рядов и итерационных процессов.

Рекомендуемая литература:

[3], [5], [7], [9], [10], [11], [13], [14], [16], [17], [19]

Цель изучения

Изучить понятие норм в векторных пространствах, ознакомиться с наиболее часто употребляемыми нормами и научиться применять их для решения прикладных задач (оценка обусловленности СЛАУ, локализация спектра матрицы, оценка невырожденности матрицы).

Следует повторить

- матрицы, операции над матрицами
- вырожденная, невырожденная, обратная матрица
- векторное пространство
- скалярное произведение, его свойства
- ортогональность
- функция модуль и ее свойства
- пространство непрерывных на $[a, b]$ функций $C[a, b]$
- понятие минимума и максимума функции
- понятие непрерывной зависимости функции от аргументов
- сходимость последовательности
- сходимость ряда
- спектр, спектральный радиус
- сингулярные числа
- след матрицы
- теорема Шура

7.1. Векторные нормы

7.1.1. Определяющие свойства векторных норм

Пусть V — векторное пространство над полем F ($\mathbb{C}$ или $\mathbb{R}$).

Функция $\|\cdot\| : V \to \mathbb{R}$ называется *векторной нормой*, если для всех $x, y \in V$ выполняются следующие условия:

1) $\|x\| \geq 0$ (неотрицательность);

1a) $\|x\| = 0 \Leftrightarrow x = 0$ (невырожденность);

2) $\|cx\| = |c| \, \|x\| \;\; \forall c \in F$ (абсолютная однородность);

3) $\|x + y\| \leq \|x\| + \|y\|$ (неравенство треугольника).

Определение 7.1. Вектор, норма которого равна единице, называется *нормированным*.

Любой ненулевой вектор x можно нормировать, умножив его на число $\lambda = \|x\|^{-1}$.

7.1.2. Примеры векторных норм

Наиболее употребительные векторные нормы на $\mathbb{C}^n$.

1*. *Евклидова норма* (l_2-норма, 2-норма)

$$\|x\|_2 \equiv \left(\sum_{j=1}^{n} |x_j|^2\right)^{\frac{1}{2}}.$$

Евклидову норму удобно применять в задачах оптимизации, т.к. она непрерывно-дифференцируема (везде, кроме нуля).

2*. *Абсолютная норма* (октаэдрическая, l_1-норма, 1-норма, Манхеттен-норма).

$$\|x\|_1 \equiv \sum_{j=1}^{n} |x_j|.$$

Абсолютная норма популярна в статистике, т.к. приводит к более робастным (устойчивым) оценкам по сравнению с классическими, основанными на регрессии.

3*. *Максимальная норма* (кубическая, l_∞-, ∞-норма).

$$\|x\|_\infty \equiv \max_{1\le j\le n} |x_j|.$$

4*. l_p-норма (*норма Гельдера* с показателем p).

Можно получить бесконечное число способов задания нормы, если в формуле, определяющей евклидову норму 1*, заменить 2 на p, $p \ge 1$:

$$\|x\|_p \equiv \left(\sum_{j=1}^{n} |x_j|^p\right)^{\frac{1}{p}}, \ p \ge 1.$$

Нормы 1* — 3* являются частным случаем нормы 4*.

Пример 7.1 (Векторные нормы в $\mathbb{C}^n$). Для вектора

$$x = \begin{bmatrix} -1 \\ -2+i \\ 3i \end{bmatrix}$$

нормы 1*— 3* равны:

$$\|x\|_2 = (|-1|^2 + |-2+i|^2 + |3i|^2)^{\frac{1}{2}} = (1+5+9)^{\frac{1}{2}} = \sqrt{15}.$$
$$\|x\|_1 = |-1| + |-2+i| + |3i| = 1 + \sqrt{5} + 3 = 4 + \sqrt{5}.$$
$$\|x\|_\infty = \max\{|-1|, |-2+i|, |3i|\} = \max\{1, \sqrt{5}, 3\} = 3.$$

Векторы

$$\frac{1}{\sqrt{15}}\begin{bmatrix} -1 \\ -2+i \\ 3i \end{bmatrix}, \frac{1}{4+\sqrt{5}}\begin{bmatrix} -1 \\ -2+i \\ 3i \end{bmatrix}, \frac{1}{3}\begin{bmatrix} -1 \\ -2+i \\ 3i \end{bmatrix}$$

нормированы по нормам 1*, 2* и 3* соответственно.

В определении векторной нормы на пространстве V не требуется конечномерности этого пространства. Пространство V может быть, например, линейным пространством $C[a,b]$ всех непрерывных вещественных или комплексных функций, заданных на вещественном отрезке $[a,b]$. Некоторые нормы на пространстве $C[a,b]$ аналогичны уже введенным векторным нормам на $\mathbb{C}^n$.

Примеры векторных норм на $C\,[a,b]^*$.

5*. L_2-норма

$$\|f\|_2 \equiv \left[\int_a^b |f(t)|^2 dt\right]^{\frac{1}{2}}.$$

6*. L_1-норма

$$\|f\|_1 \equiv \left[\int_a^b |f(t)|dt\right].$$

7*. L_∞-норма

$$\|f\|_\infty \equiv \max\{\,|f(t)|\,:\, t \in [a,b]\}\,.$$

8*. L_p-норма

$$\|f\|_p \equiv \left[\int_a^b |f(t)|^p dt\right]^{\frac{1}{p}},\ p \geq 1.$$

Пример 7.2 (Векторные нормы на $C[a,b]$).
Для функции $f(x) = \sin x,\ x \in \left[-\frac{\pi}{2}, \frac{\pi}{6}\right]$, нормы 5*—8* равны

$$\|f\|_2 = \left[\int_{-\frac{\pi}{2}}^{\frac{\pi}{6}} |\sin t|^2\, dt\right]^{\frac{1}{2}} = \left[\frac{t}{2} - \frac{\sin 2t}{4}\Bigg|_{-\frac{\pi}{2}}^{\frac{\pi}{6}}\right]^{\frac{1}{2}} = \left(\frac{\pi}{3} - \frac{\sqrt{3}}{8}\right)^{\frac{1}{2}}.$$

$$\|f\|_1 = \int_{-\frac{\pi}{2}}^{\frac{\pi}{6}} |\sin t|\, dt = 2 - \frac{\sqrt{3}}{2}.$$

$$\|f\|_\infty = \max\left\{\ |\sin x|\ :\ x \in \left[-\frac{\pi}{2}, \frac{\pi}{6}\right]\ \right\} = 1.$$

7.1.3. Алгебраические свойства векторных норм

Исходя из заданной нормы или нескольких норм можно различными способами определять новые нормы.

Теорема 7.1 (О суперпозиции векторных норм). *Если $\|\cdot\|_{\alpha_1}, \ldots, \|\cdot\|_{\alpha_m}$ — нормы на векторном пространстве V и $\|\cdot\|_\beta$ — векторная норма на $\mathbb{R}^m$, то функция $\|\cdot\| : V \to \mathbb{R}$, определенная как суперпозиция*

$$\|x\| \equiv \left\|\ \left[\ \|x\|_{\alpha_1}, \ldots, \|x\|_{\alpha_m}\right]^{\,T}\right\|_\beta,$$

является векторной нормой на V.

Упражнение 7.1. Доказать теорему 7.1, используя определение векторной нормы.

Пример 7.3 (Суперпозиция норм).

Пусть $\|\cdot\|_{\alpha_1}, \|\cdot\|_{\alpha_2}$ — некоторые нормы на векторном пространстве V, $\|\cdot\|_1$ — 1-норма на $\mathbb{R}^2$. Рассмотрим функцию

$$\|x\| \equiv \left\|\ \left[\ \|x\|_{\alpha_1}, \|x\|_{\alpha_2}\right]^{\,T}\right\|_1.$$

По свойству 1) векторной нормы и определению 2* 1-нормы функция $\|x\|$ имеет вид

$$\|x\| \equiv \|x\|_{\alpha_1} + \|x\|_{\alpha_2}$$

и согласно теореме 7.1 является векторной нормой.

Теорема 7.2 (О невырожденном преобразовании). *Если $\|\cdot\|$ — векторная норма на $\mathbb{C}^n$ и матрица $T \in \mathbb{C}^{n\times n}$ невырожденная, то функция $\|\cdot\|_T$, определенная соотношением*

$$\|x\|_T \equiv \|Tx\|\,,\ x \in \mathbb{C}^n,$$

также будет векторной нормой на $\mathbb{C}^n$.

Упражнение 7.2. Доказать теорему 7.2.

7.1.4. Аналитические свойства векторных норм

Теорема 7.3 (О непрерывности нормы). *Векторная норма* $\|\cdot\|$ *непрерывно зависит от элементов вектора, т.е. для заданного* $\varepsilon > 0\ \exists\,\delta(\varepsilon) > 0$*, такое что* $|\,\|x\| - \|y\|\,| < \varepsilon$*, как только* $|x_i - y_i| < \delta(\varepsilon)$ *для всех индексов* i*.*

Доказательство. Пусть $x, y \in \mathbb{C}^n$ — заданные векторы.

Используя единичные векторы e_i, можно представить

$$x - y = \sum_i (x_i - y_i) e_i.$$

Определим число $k = \max\limits_i \|e_i\| > 0$. Тогда справедлива цепочка соотношений:

$$\|x - y\| = \|\sum_i (x_i - y_i) e_i\| \overset{3),\,2)}{\leq} \sum_i |x_i - y_i| \cdot \|e_i\| \leq$$
$$\leq k \sum_i |x_i - y_i|.$$

Для заданного ε определим

$$\delta(\varepsilon) = \frac{\varepsilon}{kn}$$

и рассмотрим x и y, для которых $|x_i - y_i| < \delta, \quad i = \overline{1,n}$. Тогда

$$|\,\|x\| - \|y\|\,| \leq \|x - y\| \leq k \sum_i |x_i - y_i| = kn\delta = \varepsilon. \qquad \square$$

Определение 7.2. Последовательность $\left\{x^{(k)}\right\}$ векторов в V называют *сходящейся к вектору* $x \in V$ по норме $\|\cdot\|$, если

$$\left\|x^{(k)} - x\right\| \to 0 \quad \text{при} \quad k \to \infty.$$

Записывают $x^{(k)} \underset{\|\cdot\|}{\to} x$ или $\lim\limits_{k\to\infty} x^{(k)}$ по норме $\|\cdot\|$.

Заданная последовательность векторов может сходиться по одной норме и расходиться по другой.

Пример 7.4* (Сходимость по норме).

Рассмотрим последовательность $\{f_k\}$ функций в $C[0,1]$, заданных равенствами

$$f_k(x) = 0,\ 0 \leq x \leq \frac{1}{k},$$
$$f_k(x) = 2(k^{\frac{3}{2}} x - k^{\frac{1}{2}}),\ \frac{1}{k} \leq x \leq \frac{3}{2k},$$
$$f_k(x) = 2(-k^{\frac{3}{2}} x + 2k^{\frac{1}{2}}),\ \frac{3}{2k} \leq x \leq \frac{2}{k},$$
$$f_k(x) = 0,\ \frac{2}{k} \leq x \leq 1,$$

при $k = 2, 3, 4, \ldots$. Проверим сходимость к 0 последовательности $\{f_k\}$ по нормам 1,2 и ∞. Имеем

$$\|f_k\|_1 = \int\limits_0^1 |f_k(x)|\,dx = \int\limits_0^{\frac{1}{k}} |f_k(x)|\,dx + \int\limits_{\frac{1}{k}}^{\frac{3}{2k}} |f_k(x)|\,dx +$$
$$+ \int\limits_{\frac{3}{2k}}^{\frac{2}{k}} |f_k(x)|\,dx + \int\limits_{\frac{2}{k}}^{1} |f_k(x)|\,dx =$$
$$= 2k^{\frac{1}{2}} \left[\int\limits_{\frac{1}{k}}^{\frac{3}{2k}} |kx - 1|\,dx + \int\limits_{\frac{3}{2k}}^{\frac{2}{k}} |-kx + 2|\,dx \right] =$$
$$= \frac{1}{2} k^{-\frac{1}{2}} \xrightarrow[k\to\infty]{} 0,$$

$$\|f_k\|_2 = \left[\int\limits_0^{\frac{1}{k}} |f_k(x)|^2\,dx + \int\limits_{\frac{1}{k}}^{\frac{3}{2k}} |f_k(x)|^2\,dx + \right.$$

$$\left. + \int\limits_{\frac{3}{2k}}^{\frac{2}{k}} |f_k(x)|^2\,dx + \int\limits_{\frac{2}{k}}^{1} |f_k(x)|^2\,dx\right]^{\frac{1}{2}} =$$

$$= \left[\int\limits_{\frac{1}{k}}^{\frac{3}{2k}} 4k\,|kx-1|^2\,dx + \int\limits_{\frac{3}{2k}}^{\frac{2}{k}} 4k\,|-kx+2|^2\,dx\right]^{\frac{1}{2}} =$$

$$= \frac{1}{\sqrt{3}}\ \forall k,$$

$$\|f_k\|_\infty = \max\left\{0;\ 2\left(k^{\frac{3}{2}}x - k^{\frac{1}{2}}\right), x \in \left[\frac{1}{k}, \frac{3}{2k}\right];\right.$$

$$\left. 2\left(-k^{\frac{3}{2}}x + 2k^{\frac{1}{2}}\right), x \in \left[\frac{3}{2k}, \frac{2}{k}\right];\ 0\right\} =$$

$$= \max\left\{0, k^{\frac{1}{2}}, k^{\frac{1}{2}}, 0\right\} = k^{\frac{1}{2}} \xrightarrow[k\to\infty]{} \infty.$$

Таким образом, $\lim\limits_{k\to\infty} f_k = 0$ только по L_1-норме среди трех выбранных. В случае векторного пространства конечной размерности такое невозможно.

Определение 7.3. Две векторные нормы $\|\cdot\|_\alpha, \|\cdot\|_\beta$ называются *эквивалентными,* если существуют такие положительные константы c_m и c_M (зависящие от норм $\|\cdot\|_\alpha, \|\cdot\|_\beta$) , что для всех векторов $x \in V$ выполняются *соотношения эквивалентности*

$$c_m\|x\|_\alpha \le \|x\|_\beta \le c_M\|x\|_\alpha.$$

Упражнение 7.3. Доказать, что определение эквивалентности норм действительно задает отношение эквивалентности.

Теорема 7.4 (Об эквивалентности векторных норм).

Пусть $\|\cdot\|_\alpha$ и $\|\cdot\|_\beta$ — две произвольные векторные нормы в конечномерном вещественном или комплексном пространстве V. Тогда они эквивалентны.

Доказательство. Рассмотрим отношение

$$h = \frac{\|x\|_\beta}{\|x\|_\alpha}$$

на евклидовой единичной сфере

$$S = \{x \in \mathbb{R}^n(\mathbb{C}^n) : \|x\|_2 = 1\}.$$

Сфера является компактным множеством в $\mathbb{R}^n(\mathbb{C}^n)$. Так как $\|x\|_2 = 1 \ne 0$, то по свойству 1а) векторной нормы $x \ne 0$, и, значит, $\|x\|_\alpha \ne 0$. Таким образом, знаменатель не обращается в нуль и отношение непрерывно на S. По теореме Вейерштрасса у непрерывной функции h на компактном множестве S существует максимальное значение c_M и положительное (так как $h > 0$) минимальное значение c_m:

$$c_m \le h \le c_M \Rightarrow c_m \le \frac{\|x\|_\beta}{\|x\|_\alpha} \le c_M,$$

откуда следует, что соотношение теоремы верно $\forall x \in S$.

Рассмотрим теперь произвольный вектор $x \in V$. Случай $x = 0$ тривиален, поэтому считаем, что $x \ne 0$. Очевидно, что $y = \frac{x}{\|x\|_2} \in S$. Тогда по вышеуказанному

$$c_m\|y\|_\alpha \le \|y\|_\beta \le c_M\|y\|_\alpha,$$

откуда получаем

$$c_m\left\|\frac{x}{\|x\|_2}\right\|_\alpha \le \left\|\frac{x}{\|x\|_2}\right\|_\beta \le c_M\left\|\frac{x}{\|x\|_2}\right\|_\alpha,$$

откуда в силу свойства 2) векторных норм следует справедливость теоремы для $\forall x \in \mathbb{C}^n$. □

Следствие 7.1. *Если $\|\cdot\|_\alpha$ и $\|\cdot\|_\beta$ — векторные нормы в конечномерном пространстве $\mathbb{R}$ или $\mathbb{C}$ и $\left\{x^{(k)}\right\}$ — заданная последовательность векторов, то $\lim\limits_{k\to\infty} x^{(k)} = x$ по норме $\|\cdot\|_\alpha$ тогда и только тогда, когда $\lim\limits_{k\to\infty} x^{(k)} = x$ по норме $\|\cdot\|_\beta$.*

Другими словами, если две нормы эквивалентны, то любая сходящаяся по одной из этих норм последовательность сходится к тому же пределу по другой из этих норм.

В частности, покомпонентная сходимость (в любом базисе) эквивалентна сходимости по любой норме.

Доказательство следует из справедливости для любого k следующих неравенств

$$c_m \left\|x^{(k)} - x\right\|_\alpha \le \left\|x^{(k)} - x\right\|_\beta \le c_M \left\|x^{(k)} - x\right\|_\alpha,$$

т.е. $\left\|x^{(k)} - x\right\|_\alpha \to 0 \Longleftrightarrow \left\|x^{(k)} - x\right\|_\beta \to 0$.

Пример 7.5 (Константы эквивалентности). Найдем константы эквивалентности, связывающие нормы $\|\cdot\|_2$ и $\|\cdot\|_\infty$. Из неравенств

$$\max_{1\le i\le n} |x_i|^2 \le \sum_{i=1}^{n} |x_i|^2 \le n \max_{1\le i\le n} |x_i|^2$$

следует

$$\|x\|_\infty \le \|x\|_2 \le n^{\frac{1}{2}} \|x\|_\infty . \tag{7.1}$$

Для норм $\|\cdot\|_2$ и $\|\cdot\|_1$ из неравенства Коши-Буняковского имеем

$$\|x\|_1 = \sum_{i=1}^{n} |x_i| \le \left(\sum_{i=1}^{n} 1\right)^{\frac{1}{2}} \left(\sum_{i=1}^{n} |x_i|^2\right)^{\frac{1}{2}} = n^{\frac{1}{2}} \|x\|_2 .$$

Так как

$$\sum_{i=1}^{n} |x_i|^2 \le \left(\sum_{i=1}^{n} |x_i|\right)^2,$$

то

$$\|x\|_2 \le \|x\|_1 .$$

Следовательно,

$$\|x\|_2 \le \|x\|_1 \le n^{\frac{1}{2}} \|x\|_2 . \tag{7.2}$$

Заметим, что нижние неравенства в (7.1) и (7.2) достигаются на любом n-векторе x с единственной ненулевой координатой, а верхние — на любом n-векторе x, для которого $|x_1| = |x_2| = \ldots = |x_n|$. Следовательно, в (7.1) и (7.2) указаны наилучшие константы эквивалентности.

В конечномерных вещественных или комплексных векторных пространствах все нормы эквивалентны. В бесконечномерном пространстве различные нормы могут быть не эквивалентными (см. пример 7.4).

Определение 7.4*. Последовательность $\left\{x^{(k)}\right\}$ в векторном пространстве V с нормой $\|\cdot\|$ называют *последовательностью Коши,* если для любого $\varepsilon > 0$ существует натуральное число $N(\varepsilon)$ такое, что

$$\left\|x^{(k_1)} - x^{(k_2)}\right\| \le \varepsilon, \forall\, k_1, k_2 \ge N(\varepsilon).$$

Определение 7.5*. Векторное пространство V с нормой $\|\cdot\|$ называют *полным*, если каждая последовательность Коши имеет в пространстве V предел.

Конечномерные вещественные или комплексные пространства с любыми нормами являются полными. Бесконечномерные векторные пространства могут быть и неполными.

Пример 7.6* (Неполное векторное пространство).

В линейном пространстве $C\,[0,1]$ с L_1-нормой $\|f\|_1 = \int\limits_0^1 |f(t)|\,dt$ рассмотрим последовательность функций $\{f_k\}$, определенных равенствами

$$f_k(t) = 0,\ 0 \le t \le \frac{1}{2} - \frac{1}{k},$$
$$f_k(t) = \frac{k}{2}\left(t - \frac{1}{2} + \frac{1}{k}\right),\quad \frac{1}{2} - \frac{1}{k} \le t \le \frac{1}{2} + \frac{1}{k},$$
$$f_k(t) = 1,\ \frac{1}{2} + \frac{1}{k} \le t \le 1.$$

Изобразим графики функций f_{k_1}, f_{k_2} при $k_1 < k_2$:

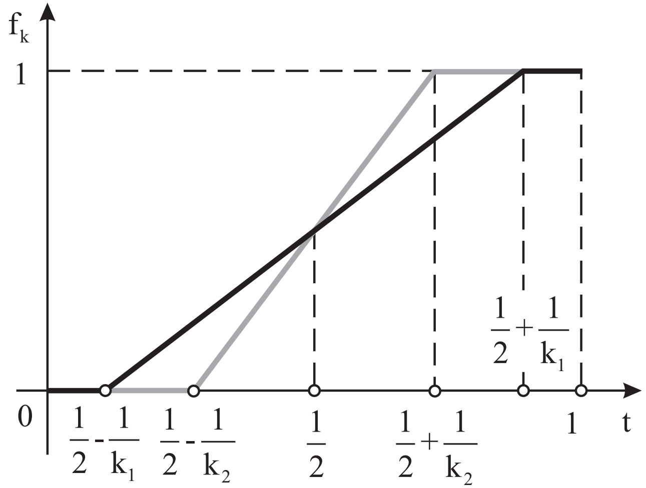


Рис. 7.1

Для k_1 и k_2 (для определенности возьмем $k_1 < k_2$) вы-

числим норму разности $\|f_{k_1} - f_{k_2}\|_1$. В силу симметричности функции $|f_{k_1} - f_{k_2}|$ относительно $t = \frac{1}{2}$ имеем:

$$\|f_{k_1} - f_{k_2}\|_1 = 2\Bigg[\int\limits_{\frac{1}{2}-\frac{1}{k_1}}^{\frac{1}{2}-\frac{1}{k_2}} \frac{k_1}{2}\left(t - \frac{1}{2} + \frac{1}{k_1}\right)dt +$$
$$+ \int\limits_{\frac{1}{2}-\frac{1}{k_2}}^{\frac{1}{2}} \left\{\frac{k_1}{2}\left(t - \frac{1}{2} + \frac{1}{k_1}\right) - \frac{k_2}{2}\left(t - \frac{1}{2} + \frac{1}{k_2}\right)\right\}dt\Bigg] =$$
$$= \frac{1}{2}\left(\frac{1}{k_1} + \frac{1}{k_2}\right).$$

При любом $\varepsilon > 0$ для любых $k_1, k_2 \ge \frac{1}{\varepsilon}$ справедливо

$$\|f_{k_1} - f_{k_2}\|_1 \le \varepsilon,$$

что означает, что последовательность $\{f_k\}$ является последовательностью Коши. Однако, как видно из рис. 7.1, не существует непрерывной на [0, 1] функции $f(t)$, такой, что $\|f_k - f\|_1 \to 0$ при $k \to \infty$, т.е. последовательность $\{f_k\}$ не является сходящейся по норме $\|\cdot\|_1$.

7.1.5. Геометрические свойства векторных норм

Если на множестве векторов определено расстояние, то с его помощью можно описать геометрические объекты, например, шары или окрестности точек в смысле этого расстояния.

Определение 7.6. *Шаром* радиуса r $(r > 0)$ с центром в точке x $(x \in V)$ называется множество

$$B_{\|\cdot\|}(r; x) \equiv \{y \in V :\ \|y - x\| \le r\}\,. \tag{7.3}$$

Определение 7.7. *Единичным шаром* для нормы $\|\cdot\|$ называется шар с $x = 0,\ r = 1$.

Шар данного радиуса с центром в точке x имеет тот же вид, что и шар такого же радиуса с центром в нуле, и может быть получен параллельным переносом последнего в точку x.

Единичный шар полностью характеризует норму.

Пример 7.7 (Единичный шар).

Множества точек единичного шара для l_1-, l_2- и l_∞- норм на $\mathbb{R}^2$ имеют вид:

$$B_{\|\cdot\|_1}(1;0) \equiv \left\{(x,y) \in \mathbb{R}^2 : \, x^2 + y^2 \le 1\right\},$$
$$B_{\|\cdot\|_2}(1;0) \equiv \left\{(x,y) \in \mathbb{R}^2 : \, |x| + |y| \le 1\right\},$$
$$B_{\|\cdot\|_\infty}(1;0) \equiv \left\{(x,y) \in \mathbb{R}^2 : \, \max\{\,|x|, |y|\} \le 1\right\}.$$

Соответствующие единичные шары изображены на рис. 7.2 a), b) и c).

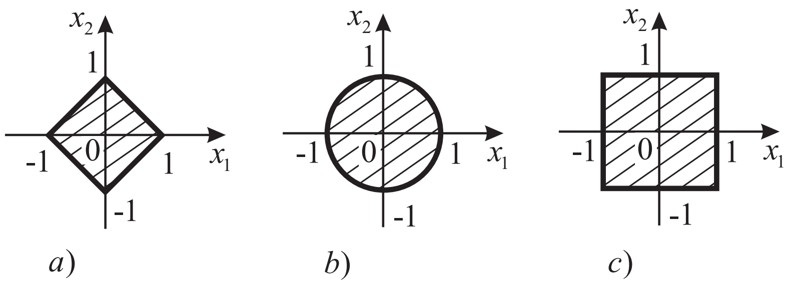


Рис. 7.2

Определение 7.8. Векторную норму называют *полиэдральной,* если ее единичный шар является многогранником. Из рис. 7.2 видно, что полиэдральными являются нормы l_1 и l_∞. В связи с этим l_1-норму называют также *октаэдрической,* а l_∞-норму — *кубической.*

Задачи и упражнения

1. Функция, для которой выполнены условия 1), 2), 3) из определения векторной нормы, но не обязательно 1a), называется *векторной полунормой.* Доказать, что для любой векторной полунормы $\|\cdot\|$ на V справедливо

   a) $|\ \|x\| - \|y\|\ | \le \|x - y\|\,,\ \forall\, x, y \in V;$

   b) $\|x - y\| - \|z - u\| \le \|x - z\| + \|y - u\|\,,\ \forall\, x, y, u, z \in V.$

2. Проверить, что нормы 1*— 4* действительно являются векторными нормами на $\mathbb{C}^n$.

3. Пусть $p_1, p_2, \ldots, p_n$ — положительные числа. Являются ли векторными нормами на $\mathbb{C}^n$ следующие выражения?

   a) $\|x\| = \sum\limits_{i=1}^{n} p_i\,|x_i|$; b) $\|x\| = \left(\sum\limits_{i=1}^{n} p_i\,|x_i|^2\right)^{\frac{1}{2}}$;

   c) $\|x\| = \max\{p_1\,|x_1|\,,..,p_n\,|x_n|\}$; d) $\|x\| = \min\{|x_1|\,,..,|x_n|\}$;

   e) $\|x\| = \max\limits_{t\in[0,1]}\left\|\sum\limits_{i=1}^{n} x_i t^{i-1}\right\|$; f) $\|x\| = \max\limits_{t\in[0,1]}\left\|\sum\limits_{i=1}^{n} x_i t^{i}\right\|$.

4*. Показать, что $\|x\|_\infty = \lim\limits_{p\to\infty}\|x\|_p\,,\ \forall x \in \mathbb{C}^n$.

5*. Пусть задана точка $x_0 \in [a, b]$. Показать, что функционал $\|f\|_{x_0} \equiv |f(x_o)|$ на $C\,[a, b]$ является полунормой, но не нормой.

6. Проверить, что сумма конечного числа векторных норм (полунорм) является векторной нормой (полунормой).

7. Проверить, что произведение векторной нормы (полунормы) на положительное число является векторной нормой (полунормой).

8. Проверить, что для векторных норм $\|\cdot\|_\alpha$ и $\|\cdot\|_\beta$ функция $\|\cdot\|$, определенная равенством $\|x\| \equiv \max\{\,\|x\|_\alpha\,, \|x\|_\beta\}$, является нормой.

9. Пользуясь нормами 1*— 4* и теоремой 7.1, сконструировать новые векторные нормы.

10. Пользуясь теоремами 7.1 и 7.2 (без вычислений!), показать, что функция $\|x\| \equiv \left(|2x_1 - 3x_2|^2 + |x_2|^2\right)^{\frac{1}{2}}$ является нормой на $\mathbb{C}^2$.

11. Найти константы эквивалентности, связывающие нормы $\|\cdot\|_\infty$ и $\|\cdot\|_1$, а также векторы, на которых они достигаются.

12. Записать определение единичного шара для произвольной векторной нормы $\|\cdot\|$.

13. Показать, что для любых $r > 0$ и $x \in V$ справедливы представления $B(r;x) = \{y + x : y \in B(r;0)\} = x + B(r;0)$.

14. Какие включения можно установить для единичных шаров l_p-норм при различных p $(p \geq 1)$?

15. Изобразить единичные шары для l_3-, l_6-норм на $\mathbb{C}^2$.

16. Пусть $\|\cdot\|_\alpha$ и $\|\cdot\|_\beta$ — две нормы на векторном пространстве V, $B_{\|\cdot\|_\alpha}, B_{\|\cdot\|_\beta}$ — соответственные единичные шары. Показать, что

$$\|x\|_\alpha \leq \|x\|_\beta \ \forall x \in V \Leftrightarrow B_{\|\cdot\|_\beta} \subset B_{\|\cdot\|_\alpha} \Rightarrow \|x\|_p \leq \|x\|_{p-1} \ \forall p.$$

17. Что происходит с единичным шаром при умножении нормы на положительную константу C?

18. Найти l_1-, l_2- и l_∞- нормы для векторов

$$x = (3, 4, -3i, 1)^T, y = (2, 1, -4, 2)^T, z = (1 + i, 1 - i, 1, 4i)^T,$$
$$u = (0, -1, 2, -2, 4)^T, v = (i, 2, 1 - i, 0, 1 + i)^T.$$

19. Доказать, что если функция $\langle\cdot,\cdot\rangle$ является скалярным произведением на V, то функция $\|x\| = \langle x, x\rangle^{\frac{1}{2}}$ задает векторную норму на V. При этом говорят, что норма $\|\cdot\|$ *порождена скалярным произведением* $\langle\cdot,\cdot\rangle$.

20. Векторная норма порождена каким-нибудь скалярным произведением тогда и только тогда, когда имеет место *тождество параллелограмма*

$$\frac{1}{2}\left(\|x + y\|^2 + \|x - y\|^2\right) = \|x\|^2 + \|y\|^2. \tag{7.4}$$

Доказать, что евклидова норма порождена скалярным произведением.

21*. Показать, что функция $\|x\|_\infty \equiv \max\limits_{1\leq i\leq n} |x_i|$ на $\mathbb{C}^n$ является векторной нормой, которая не порождается скалярным произведением.

22. Доказать, что если концы отрезка принадлежат единичному шару, то ему принадлежит и весь отрезок.

23*. Пусть D — положительно (неотрицательно) определенная матрица и определена функция $||x||_D = \sqrt{x^T D x}$, $x \in \mathbb{C}^n$. Доказать, что функция $||x||_D$ является векторной нормой.

Вопросы для повторения

1. Что такое векторная норма?
2*. Что такое векторная полунорма?
3*. Какая норма называется нормой, порожденной скалярным произведением?
4. Какой вектор называется нормированным?
5. Как нормировать заданный вектор?
6. Что такое единичный шар нормы $\|\cdot\|$?
7. Какие точки принадлежат одновременно границам единичных шаров всех l_p-норм на $\mathbb{R}^2$?
8. Какие из l_p-норм являются полиэдральными? Проиллюстрировать ответ геометрическими построениями.
9. Используя евклидову норму, описать
— единичный шар в $\mathbb{R}^n$;
— шар с центром в точке $c = (c_1, \ldots, c_n)^T$ радиуса r.
10. Объяснить, почему равенство $\|x - y\| = \|y - x\|$ верно для любой нормы.

7.2. Матричные нормы

Во многих случаях бывает полезно иметь возможность приписать матрице некоторое однозначно определенное неотрицательное число в качестве меры ее отклонения от нулевой матрицы. Для комплексного числа такая мера определяется его модулем. Для векторов из $\mathbb{C}^n$ величина может быть измерена нормой.

Здесь вводится понятие нормы матрицы, которая служит ее мерой, показывается, что существует много таких норм, устанавливается связь матричных и векторных норм, описываются некоторые приложения матричных норм.

7.2.1. Определение и примеры матричных норм

Определение 7.9. Функция $\|\cdot\| : \mathbb{C}^{n\times m} \to \mathbb{R}$ называется *матричной нормой,* если для любой $A \in \mathbb{C}^{n\times m}$ выполняются следующие условия:

1) $\|A\| \geq 0$ (неотрицательность);

1a) $\|A\| = 0 \Leftrightarrow A = 0$ (невырожденность);

2) $\|\alpha A\| = |\alpha|\,\|A\| \ \ \forall \alpha \in \mathbb{C}$ (абсолютная однородность);

3) $\|A + B\| \leq \|A\| + \|B\|$ (неравенство треугольника);

4) $\|AB\| \leq \|A\| \cdot \|B\|$ (кольцевое свойство)

для любых матриц A, B, для которых соответствующие операции имеют смысл.

Определение 7.10. Если норма удовлетворяет всем аксиомам, кроме 4), то такую норму называют *аддитивной* или *обобщенной матричной нормой.* Если удовлетворяются все аксиомы 1) — 4), то такая норма матрицы называется *мультипликативной.*

Всякую аддитивную норму умножением на достаточно большую положительную константу можно превратить в мультипликативную. Поэтому везде, где не оговорено противное, под матричной нормой будем понимать мультипликативную матричную норму.

Также, как и векторная норма, обобщенная матричная норма является непрерывной функцией от матричного аргумента, поэтому «малые» изменения в элементах матрицы приводят к «малому» изменению в матричной норме.

Наиболее употребительные матричные нормы на $\mathbb{C}^{n\times m}$.

1*. *Столбцовая норма* (1-норма)

$$\|A\|_1 \equiv \max_{1\leq j\leq m} \sum_{i=1}^{n} |a_{ij}|.$$

2*. *Строчная норма* (∞-норма)

$$\|A\|_\infty \equiv \max_{1\leq i\leq n} \sum_{j=1}^{m} |a_{ij}|.$$

3*. *Евклидова норма* (норма Фробениуса, норма Шура, норма Гильберта—Шмидта, сферическая норма)

$$\|A\|_E \equiv \|A\|_F \equiv \Bigl(\sum_{i=1}^{n}\sum_{j=1}^{m} |a_{ij}|^2\Bigr)^{\frac{1}{2}}.$$

4*. *Спектральная норма* (вторая норма, норма Гильберта)

$$\|A\|_2 \equiv \max_{1\leq j\leq n} \sigma_j,$$

где σ_j — сингулярные числа матрицы A.

5*. M-норма

$$\|A\|_M \equiv (nm)^{\frac{1}{2}} \max_{i,j} |a_{ij}| \ .$$

6*. l_1-норма

$$\|A\|_{l_1} \equiv \sum_{i=1}^{n} \sum_{j=1}^{m} |a_{ij}|.$$

Пример 7.8 (Евклидова норма).

Покажем, что евклидова норма 3* удовлетворяет всем аксиомам матричной нормы.

Аксиомы 1) — 2) легко проверить, используя свойства функции модуль. Проверим аксиому 3). При этом будем использовать *неравенство Коши—Шварца:*

$$\sum_{i=1}^{n} x_i y_i \le \Big(\sum_{i=1}^{n} x_i^2\Big)^{\frac{1}{2}} \Big(\sum_{i=1}^{n} y_i^2\Big)^{\frac{1}{2}}.$$

Имеем:

$$\|A+B\|_E^2 = \sum_{i=1}^{n} \sum_{j=1}^{m} |a_{ij}+b_{ij}|^2 = \sum_{i=1}^{n} \sum_{j=1}^{m} (a_{ij}+b_{ij}) \left(\overline{a}_{ij}+\overline{b}_{ij}\right) =$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{m} \left[|a_{ij}|^2 + |b_{ij}|^2 + 2Re\left(a_{ij}\overline{b}_{ij}\right)\right] \le$$

$$\le \sum_{i=1}^{n} \sum_{j=1}^{m} \left[|a_{ij}|^2 + |b_{ij}|^2 + 2\left|a_{ij}\overline{b}_{ij}\right|\right] \le$$

$$\le \sum_{i=1}^{n} \sum_{j=1}^{m} \left[|a_{ij}|^2 + |b_{ij}|^2\right] +$$

$$+ 2\left[\sum_{i=1}^{n} \sum_{j=1}^{m} |a_{ij}|^2 \sum_{i=1}^{n} \sum_{j=1}^{m} |b_{ij}|^2\right]^{\frac{1}{2}} =$$

$$= (\|A\|_E + \|B\|_E)^2 .$$

Проверим аксиому 4) ($B \in \mathbb{C}^{n\times r}$):

$$\|AB\|_E^2 = \sum_{i=1}^{n} \sum_{j=1}^{m} \left|\sum_{k=1}^{m} a_{ik} b_{kj}\right|^2 \le$$

$$\le \sum_{i=1}^{n} \sum_{j=1}^{r} \Big(\sum_{k=1}^{m} |a_{ik}|^2 \sum_{s=1}^{m} |b_{sj}|^2\Big) =$$

$$= \sum_{i=1}^{n} \sum_{k=1}^{m} |a_{ik}|^2 \sum_{s=1}^{m} \sum_{j=1}^{r} |b_{sj}|^2 = \|A\|_E^2 \, \|B\|_E^2 \, .$$

Теорема 7.5. *Для евклидовой нормы матрицы справедливы представления*

$$\|A\|_E^2 = \mathrm{tr} AA^* = \mathrm{tr} A^* A; \tag{7.5}$$

$$\|A\|_E^2 = \sum_{i} \sigma_i^2, \tag{7.6}$$

где σ_i — сингулярные числа матрицы A.

Доказательство. Представление (7.5) легко следует из леммы 2.5 и определения евклидовой нормы.

Из определения сингулярных чисел и теоремы 5.6 имеем

$$\mathrm{tr} A^* A = \sum_{i} \lambda_i (A^* A) = \sum_{i} \sigma_i^2, \tag{7.7}$$

что и завершает доказательство теоремы. □

Пример 7.9 (Матричные нормы).

Для матрицы

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}$$

нормы 1^*-3^* равны:

$$\begin{aligned} \|A\|_1 &= \max\{\ |1|+|3|+|5|\,,\ \ |2|+|4|+|6|\ \} = 12,\\ \|A\|_\infty &= \max\{\ |1|+|2|\,,\ \ |3|+|4|\,,\ \ |5|+|6|\ \} = 11,\\ \|A\|_E &= \left(1^2+2^2+3^2+4^2+5^2+6^2\right)^{\frac{1}{2}} = \sqrt{91}. \end{aligned}$$

Для вычисления спектральной нормы 4^* найдем сингулярные числа матрицы A. Воспользуемся тем, что ненулевые собственные значения матриц A^*A и AA^* всегда совпадают и найдем собственные значения матрицы

$$A^*A = \begin{bmatrix} 35 & 44\\ 44 & 56 \end{bmatrix}.$$

Имеем

$$\lambda_1\left(A^*A\right) = \frac{91-\sqrt{8185}}{2},\quad \lambda_2\left(A^*A\right) = \frac{91+\sqrt{8185}}{2},$$

откуда $\|A\|_2 = \left(\frac{91+\sqrt{8185}}{2}\right)^{\frac{1}{2}} \approx 9.53$.

Нормы 5^*-6^* для матрицы A равны

$$\begin{aligned} \|A\|_M &= (3\times 2)^{\frac{1}{2}}\max\{\ |1|\,,|2|\,,|3|\,,\ |4|\,,|5|\,,|6|\ \} = 6\sqrt{6},\\ \|A\|_{l_1} &= |1|+|2|+\ |3|+|4|+\ |5|+|6|\ = 21. \end{aligned}$$

Так же как и векторные нормы в конечномерных пространствах, все обобщенные матричные нормы на $\mathbb{C}^{m\times n}$ эквивалентны. Поэтому для любых двух матричных норм $\|\cdot\|_\alpha$ и $\|\cdot\|_\beta$ существует положительная константа c_M, зависящая от норм, такая, что

$$\|A\|_\alpha \le c_M\,\|A\|_\beta\,. \tag{7.8}$$

Для норм 1^*-6^* эти константы указаны в следующей таблице:

Таблица 7.1 – Константы эквивалентности матричных норм

$\alpha\backslash\beta$	**1**	**2**	∞	l_1	**E**	**M**
1	1	$\sqrt{n}$	n	1	$\sqrt{n}$	1
2	$\sqrt{n}$	1	$\sqrt{n}$	1	1	1
∞	n	$\sqrt{n}$	1	1	$\sqrt{n}$	1
l_1	n	$n^{\frac{3}{2}}$	n	1	n	n
E	$\sqrt{n}$	$\sqrt{n}$	$\sqrt{n}$	1	1	1
M	n	n	n	n	n	1

7.2.2. Связь матричных и векторных норм

Определение 7.11. Матричная норма $\|\cdot\|_M$ называется *согласованной* с векторной нормой $\|\cdot\|_V$, если

$$\|Ax\|_V \le \|A\|_M\,\|x\|_V \tag{7.9}$$

для любой матрицы A и всех векторов x.

Определение 7.12. Пусть задана векторная норма $\|\cdot\|_V$. Тогда числовая функция

$$\|A\| \equiv \max_{\|x\|_V\neq 0}\frac{\|Ax\|_V}{\|x\|_V} = \max_{\|x\|_V=1}\|Ax\|_V \tag{7.10}$$

называется нормой матрицы, *подчиненной (индуцированной)* векторной норме $\|\cdot\|_V$.

Теорема 7.6 (Подчиненная векторная норма). *Функция (7.10) определена и является согласованной нормой в $\mathbb{C}^{m\times n}$ для любых норм в $\mathbb{C}^m$ и $\mathbb{C}^n$.*

Доказательство. Во-первых заметим, что максимум в (7.10) достигается, поскольку функция $\|Ax\|$ непрерывно зависит от

x и единичная сфера $\|x\|_V = 1$ является компактным множеством.

Первые три условия из определения матричной нормы легко следуют из аналогичных условий для векторной нормы:

1) $\|A\| \geq 0 \Leftrightarrow \max\limits_{\|x\|_V \neq 0} \frac{\|Ax\|_V}{\|x\|_V} \geq 0$;

1a)

$\|A\| = 0 \Leftrightarrow \max \frac{\|Ax\|}{\|x\|} = 0 \Leftrightarrow \|Ax\| = 0 \, \forall x \in \mathbb{C}^n \Leftrightarrow Ax = 0 \Leftrightarrow A = 0$;

2)

$\|\alpha A\| = \max\limits_{\|x\|=1} \|\alpha Ax\| = \max\limits_{\|x\|=1}\big(|\alpha| \, \|Ax\|\big) = |\alpha| \max\limits_{\|x\|=1} \|Ax\| = |\alpha| \cdot \|A\|$;

3)

$$\|A+B\| = \max_{\|x\|=1} \|(A+B)x\| = \max_{\|x\|=1} \|Ax + Bx\| \leq$$
$$\leq \max_{\|x\|=1} \|Ax\| + \max_{\|x\|=1} \|Bx\| = \|A\| + \|B\|.$$

Таким образом, функция (7.10) является матричной нормой. Докажем, что для неё выполняется условие согласованности (7.9). По определению максимума

$$\frac{\|Ax\|_V}{\|x\|_V} \leq \max_{\|x\|_V \neq 0} \frac{\|Ax\|_V}{\|x\|_V} = \|A\|_M,$$

следовательно,

$$\|Ax\|_V \leq \|A\|_M \cdot \|x\|_V, \quad \forall x \in \mathbb{C}^n,$$

значит, согласно (7.9) матричная норма $\|\cdot\|_M$, подчиненная векторной норме $\|\cdot\|_V$, согласована с этой нормой.

4) Покажем, наконец, что для подчиненной нормы выполняется также кольцевое свойство. Рассмотрим (7.9) для AB:

$$\|AB\| = \max_{\|x\|=1} \|(AB)x\| = \max_{\|x\|=1} \|A(Bx)\| \overset{(7.9)}{\leq}$$
$$\leq \max_{\|x\|=1} (\|A\| \ \|Bx\|) = \|A\| \max_{\|x\|=1} \|Bx\| = \|A\| \, \|B\|,$$

откуда следует, что для функции (7.10) удовлетворяется и четвертое (кольцевое) условие. □

Теорема 7.7. *Для подчиненной матричной нормы* $\|\cdot\|$ *верно*

$$\|Ax\| \leq \|A\| \cdot \|x\|, \tag{7.11}$$

$$\|E\| = 1. \tag{7.12}$$

Доказательство. Соотношение (7.11) доказано в теореме 7.6. Далее,

$$\|E\| = \max_{\|x\|=1} \|Ex\| = \max_{\|x\|=1} \|x\| = 1.$$ □

Упражнение 7.4. Используя аксиому невырожденности матричной нормы, определение обратной матрицы и кольцевое свойство, доказать, что для любой матричной нормы $\|\cdot\|$ верно

$$\|E_n\| \geq 1; \tag{7.13}$$

$$\|A^{-1}\| \geq \frac{1}{\|A\|}. \tag{7.14}$$

Пример 7.10 (Норма, подчиненная кубической).

Найдем матричную норму, подчиненную векторной норме $\|\cdot\|_\infty$. Получим оценку сверху для величины $\|Ax\|_\infty$:

$$\|Ax\|_\infty = \max_i \Big|\sum_j a_{ij} x_j\Big| \leq \max_i \Big(\sum_j |a_{ij}| \max_j |x_j|\Big) \leq$$
$$\leq \max_i \Big(\sum_j |a_{ij}|\Big) \, \|x\|_\infty = \|A\|_\infty \|x\|_\infty \Rightarrow$$
$$\Rightarrow \|A\|_\infty \geq \frac{\|Ax\|_\infty}{\|x\|_\infty}.$$

Покажем, что эта оценка достигается. Пусть максимум по i имеет место при $i = l$. Тогда возьмем

$$x = \Big(\mathrm{sign}(a_{l1}), \ldots, \mathrm{sign}(a_{l\,n})\Big).$$

Имеем $\|x\|_\infty = 1$ и точные равенства во всей цепочке выше. Таким образом,

$$\|A\|_\infty = \max_i \Big(\sum_j |a_{ij}|\Big)$$

— подчиненная векторной норме $\|\cdot\|_\infty$.

Пример 7.11 (Норма, подчиненная евклидовой).

Найдем матричную норму, подчиненную векторной норме $\|\cdot\|_2$. По определению матричной нормы, подчиненной евклидовой векторной норме, имеем:

$$\|A\|_2 = \max_{\|x\|_2 \neq 0} \frac{\|Ax\|_2}{\|x\|_2} = \max_{\|x\|_2 \neq 0} \sqrt{\frac{\langle Ax, Ax\rangle}{\langle x, x\rangle}} = \max_{\|x\|_2 \neq 0} \sqrt{\frac{\langle A^*Ax, x\rangle}{\langle x, x\rangle}}.$$

Т.о., имеем

$$\max_{\|x\|_2 \neq 0} \sqrt{\frac{\langle A^*Ax, x\rangle}{\langle x, x\rangle}} = \max_i \sqrt{\lambda_i(A^*A)} = \max_i \sigma_i(A) = \|A\|_2\,,$$

а равенство достигается на соответствующем собственном векторе.

Таким образом, спектральная матричная норма $\|\cdot\|_2$ подчинена евклидовой векторной норме $\|\cdot\|_2$.

Упражнение 7.5. Показать, что столбцовая матричная норма $\|\cdot\|_1$ подчинена октаэдрической векторной норме $\|\cdot\|_1$.

Среди всех матричных норм, согласованных с заданной векторной нормой, подчиненная норма является *минимальной* в том смысле, что в неравенстве $\|Ax\|_V \le \|A\|_M \|x\|_V$ число $\|A\|_M$ нельзя уменьшить.

Таблица 7.2 – Подчиненные нормы

Векторная норма		**Подчиненная матричная норма**	
октаэдрическая	$\|x\|_1$	столбцовая	$\|A\|_1$
евклидова	$\|x\|_2$	спектральная	$\|A\|_2$
кубическая	$\|x\|_\infty$	строчная	$\|A\|_\infty$

7.2.3. Приложения матричных норм

Рассматриваются вопросы применения матричных норм для оценки собственных значений, анализа обратимости матриц, оценивания ошибок при вычислении обратной матрицы и решении линейных алгебраических уравнений.

Определение 7.13. *Спектром* матрицы $A \in \mathbb{C}^{n\times n}$ (обозначается $\sigma(A)$) называется совокупность всех собственных значений $\lambda(A) \in \mathbb{C}$ матрицы A. *Спектральным радиусом* матрицы A называется неотрицательное вещественное число

$$\rho(A) = \max\{\, |\lambda| : \lambda \in \sigma(A)\}.$$

Одна важная область использования матричных норм — определение границ спектра матрицы.

Лемма 7.1. *Для любой матричной нормы $\|\cdot\|$ и любой матрицы $A \in \mathbb{C}^{n\times n}$ выполняется неравенство*

$$\rho(A) \le \|A\|\,. \tag{7.15}$$

Доказательство. Действительно, возьмем векторную норму, согласованную с заданной матричной. Тогда из $Ax = \lambda x$ для собственных векторов x имеем

$$|\lambda|\,\|x\| \le \|A\|\cdot\|x\| \;\Rightarrow\; |\lambda| \le \|A\|\,.$$

Отсюда с учетом определения спектрального радиуса вытекает справедливость (7.11). $\square$

Лемма 7.2. *Пусть $A \in \mathbb{C}^{n\times n}$ и задано число $\varepsilon > 0$. Существует по крайней мере одна матричная норма $\|\cdot\|$, для которой имеют место оценки*

$$\rho(A) \le \|A\| \le \rho(A) + \varepsilon.$$

Доказательство.* В силу теоремы Шура об унитарной триангуляризации найдутся такая унитарная матрица U и верхняя треугольная матрица T, что $A = UTU^*$. Положим $D_t \equiv diag\left[t, t^2, \ldots, t^n\right]$ и вычислим

$$D_t T D_t^{-1} = \begin{bmatrix} \lambda_1 & t^{-1}d_{12} & t^{-2}d_{13} & \ldots & t^{-n+1}d_{1n} \\ 0 & \lambda_2 & t^{-1}d_{23} & \ldots & t^{-n+2}d_{2n} \\ 0 & 0 & \lambda_3 & \ldots & t^{-n+3}d_{3n} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & t^{-1}d_{n-1,n} \\ 0 & 0 & 0 & \ldots & \lambda_n \end{bmatrix}.$$

Таким образом, при достаточно большом t сумма модулей наддиагональных элементов матрицы $D_t T D_t^{-1}$ не будет превосходить ε, в частности, $\left\|D_t T D_t^{-1}\right\|_1 \le \rho(A) + \varepsilon$. Определим матричную норму $\|\cdot\|$ согласно формуле

$$\|B\| = \left\|D_t U^* B U D_t^{-1}\right\|_1.$$

Таким образом, выбирая t достаточно большим, получаем матричную норму, для которой $\|A\| \le \rho(A) + \varepsilon$. Так как неравенство $\rho(A) \le \|A\|$ верно для любой нормы, то утверждение леммы полностью доказано. □

Таким образом, спектральный радиус $\rho(A)$ каждой фиксированной матрицы $A \in \mathbb{C}^{n\times n}$ — это точная нижняя грань значений всех матричных норм матрицы A.

Во многих приложениях, например при анализе итерационных процессов, важна сходимость к 0 последовательности A^k, $k = 1, 2, \ldots$.

Определение 7.14. Матрица $A \in \mathbb{C}^{n\times n}$ называется *сходящейся*, если

$$\lim_{k\to\infty} A^k = 0.$$

Описание сходящихся матриц можно получить, используя оценку спектрального радиуса.

Лемма 7.3. *Пусть задана матрица $A \in \mathbb{C}^{n\times n}$. Если существует матричная норма $\|\cdot\|$, для которой $\|A\| < 1$, то* $\lim\limits_{k\to\infty} A^k = 0$.

Доказательство. Если $\|A\| < 1$, то с учетом кольцевого свойства 5)

$$\left\|A^k\right\| \le \|A\|^k \underset{k\to\infty}{\longrightarrow} 0,$$

т.е.

$$A^k \underset{\|\cdot\|}{\longrightarrow} 0.$$

В силу эквивалентности всех норм на $\mathbb{R}^{n^2}$ имеем $A^k \to 0$ по норме $\|\cdot\|_\infty$, т.е. поэлементно. □

Теорема 7.8. *Матрица $A \in \mathbb{C}^{n\times n}$ является сходящейся тогда и только тогда, когда*

$$\rho(A) < 1. \tag{7.16}$$

Доказательство. Пусть A— сходящаяся, т.е. $A^k \to 0$. Рассмотрим собственный вектор $x \ne 0:\ Ax = \lambda x$. Тогда

$$A^k x = \lambda^k x \to 0,$$

что возможно только при $|\lambda| < 1$. Так как это неравенство выполняется для $\forall \lambda(A)$, то $\rho(A) < 1$.

В обратную сторону. Если $\rho(A) < 1$, то по лемме 7.2 $\exists \|\cdot\|:$ $\|A\| < 1$. Тогда по лемме 7.3 $A^k \to 0$ при $k \to \infty$. □

Следствие 7.2. *Для любой матрицы $A \in \mathbb{C}^{n\times n}$ и любой матричной нормы*

$$\rho(A) = \lim_{k\to\infty} \left\|A^k\right\|^{\frac{1}{k}}.$$

Доказательство. Из соотношений

$$\rho(A)^k = \rho(A^k) \le \left\|A^k\right\|$$

вытекает неравенство

$$\rho(A) \le \left\|A^k\right\|^{\frac{1}{k}} \ \ \forall\, k = 1, 2, \ldots. \tag{7.17}$$

Пусть задано $\varepsilon > 0$. Рассмотрим матрицу

$$\tilde{A} \equiv [\rho(A) + \varepsilon]^{-1} A.$$

Ее спектральный радиус $\rho(\tilde{A}) < 1$, значит, эта матрица сходящаяся. Таким образом, $\left\|\tilde{A}^k\right\| \to 0$ при $k \to \infty$ и существует такой номер $N(A, \varepsilon)$, что неравенство $\left\|\tilde{A}^k\right\| < 1$ верно для всех степеней $k \ge N$. Это означает, что $\left\|A^k\right\| \le [\rho(A) + \varepsilon]^k$ для всех $k \ge N$, или, эквивалентно, $\left\|A^k\right\|^{\frac{1}{k}} \le \rho(A) + \varepsilon$ для всех $k \ge N$. Поскольку величина ε произвольна, вспоминая неравенство (7.17), приходим к заключению, что предел $\lim\limits_{k\to\infty} \left\|A^k\right\|^{\frac{1}{k}}$ существует и равен $\rho(A)$. □

Матричную норму можно использовать для анализа обратимости матрицы.

Теорема 7.9. *Матрица $A \in \mathbb{C}^{n\times n}$ обратима, если существует такая матричная норма $\|\cdot\|$, что $\|E_n - A\| < 1$. При этом условии*

$$A^{-1} = \sum_{k=0}^{\infty} (E - A)^k. \tag{7.18}$$

Доказательство. Пусть $\|E_n - A\| < 1$. Поскольку радиус сходимости числового ряда $\sum\limits_{k=0}^{\infty} z^k$ равен 1, то в этом случае ряд $\sum\limits_{k=0}^{\infty} \|E - A\|^k$ сходится абсолютно. Тогда по критерию Коши

$$\sum_{s=n+1}^{n+k} \|E - A\|^s \underset{n\to\infty}{\longrightarrow} 0. \tag{7.19}$$

Далее, поскольку из кольцевого свойства нормы следует:

$$\left\|\sum_{s=n+1}^{n+k} (E - A)^s\right\| \le \sum_{s=n+1}^{n+k} \|E - A\|^s,$$

то с учетом (7.19) имеем

$$\left\|\sum_{s=n+1}^{n+k} (E - A)^s\right\| \underset{n\to\infty}{\longrightarrow} 0,$$

откуда следует сходимость матричного ряда $\sum\limits_{s=0}^{\infty} (E - A)^s$ по норме $\|\cdot\|$. Из эквивалентности матричных норм следует покомпонентная сходимость, т.е. элементы $\sum\limits_{s=0}^{\infty} [(E - A)^s]_{ij}$ сходятся к некоторым элементам b_{ij}.

Покажем, что $B = [b_{ij}] = A^{-1}$. Для этого умножим обе части равенства (7.18) на A :

$$\begin{aligned}
A \cdot B &= A(E + (E - A) + (E - A)^2 + \ldots) = \\
&= [E - (E - A)]\left[E + (E - A) + (E - A)^2 + \ldots\right] = \\
&= E + (E - A) + + \cdots - (E - A) - \cdots = E,
\end{aligned}$$

что согласно определению обратной матрицы и завершает доказательство. □

На основании понятия нормы вводится понятие числа обусловленности матрицы, полезное для оценивания ошибок, возникающих при вычислении обратной матрицы и при решении систем линейных алгебраических уравнений численными методами.

Определение 7.15. Величина

$$\operatorname{cond}(A) \equiv \begin{cases} \left\|A^{-1}\right\| \cdot \|A\|, & \text{если } A \text{ невырождена}, \\ \infty, & \text{если } A \text{ вырождена}. \end{cases} \tag{7.20}$$

называется *числом обусловленности* матрицы A по отношению к матричной норме $\|\cdot\|$.

Свойства числа обусловленности

1^0. Для любой матричной нормы $\operatorname{cond}(A) \geq 1$.

2^0. Для любой матрицы A и любого числа α выполняется равенство $\operatorname{cond}(A) = \operatorname{cond}(\alpha A)$.

3^0. Для спектральной и евклидовой норм число обусловленности не меняется при умножении матрицы слева и справа на любую унитарную матрицу.

4^0. Для 1-, 2-, ∞-норм и евклидовой нормы число обусловленности не меняется при перестановке строк и столбцов.

5^0. Для спектральной нормы число обусловленности любой матрицы равно отношению максимального сингулярного числа к минимальному.

Доказательство. 1^0. Легко следует из определения числа обусловленности и (7.14).

2^0. Вытекает непосредственно из определения числа обусловленности и свойства абсолютной однородности матричной нормы.

3^0. Пусть $A \in \mathbb{C}^{n\times n}$ — заданная матрица, $U \in \mathbb{C}^{n\times n}$ — унитарная матрица. Во-первых покажем, что спектральная и евклидова нормы унитарно инвариантны, т.е. не меняются при умножении матрицы на унитарную. Воспользуемся теоремой 7.5, определениями унитарной матрицы и сингулярных чисел:

$$\|UA\|_E^2 = \operatorname{tr}[(UA)^*UA] = \operatorname{tr} A^*U^*UA = \operatorname{tr} A^*A = \|A\|_E^2;$$
$$\|UA\|_2^2 = \max_i \sigma_i^2(UA) = \max_i \lambda_i[(UA)^*(UA)] =$$
$$= \max_i \lambda_i[A^*A] = \max_i \sigma_i^2(A) = \|A\|_2^2.$$

Умножение справа на унитарную матрицу рассматривается аналогично. Далее согласно положениям b) и d) теоремы 5.2 матрица U^{-1} также унитарна, поэтому верно

$$\|(UA)^{-1}\|_E = \|A^{-1}U^{-1}\|_E = \|A^{-1}\|_E, \ \|(UA)^{-1}\|_2 = \|A^{-1}\|_2.$$

Доказательство далее следует из определения числа обусловленности.

4^0. Перестановка строк (столбцов) матрицы эквивалентна умножению ее слева (справа) на матрицу перестановки, которая является унитарной. Таким образом, для спектральной и евклидовой норм утверждение 4^0 следует из 3^0. Для 1- и ∞-норм результат очевиден из определения этих норм.

5^0. Из определений числа обусловленности, спектральной нормы и свойства собственных значений $\lambda_i(A^{-1}) = \lambda_i^{-1}(A)$ имеем:

$$\operatorname{cond}_2 A = \|A\|_2 \cdot \|A^{-1}\|_2 = \max_i \sigma_i(A) \max_i \sigma_i(A^{-1}) =$$
$$= \max_i (\lambda_i[A^*A])^{\frac{1}{2}} \max_i (\lambda_i[(A^*A)^{-1}])^{\frac{1}{2}} =$$
$$= \frac{\max_i (\lambda_i[A^*A])^{\frac{1}{2}}}{\min_i (\lambda_i(A^*A])^{\frac{1}{2}}} = \frac{\max_i \sigma_i(A)}{\min_i \sigma_i(A)}. \qquad \square$$

Упражнение 7.6. Доказать подробно свойства числа обусловленности.

* Приложение числа обусловленности к оценке точности вычисления обратной матрицы

Во многих алгоритмах эффект ошибок округления при вычислениях можно смоделировать при помощи возмущений лишь начальных данных.

Пусть задана невырожденная матрица $A \in \mathbb{C}^{n\times n}$ и пусть обратная к ней матрица A^{-1} вычисляется с некоторой погрешностью (например, в силу ошибок округления) таким образом, что в действительности вместо точной матрицы A^{-1} имеем матрицу $(A+\delta A)^{-1}$, где возмущение $\delta A \in \mathbb{C}^{n\times n}$ достаточно «мало», так что матрица $A+\delta A$ обратима.

Оценим ошибку $A^{-1}-(A+\delta A)^{-1}$, совершаемую при вычислении обратной матрицы.

Теорема 7.10. *Пусть* $A, \delta A \in \mathbb{C}^{n\times n}$ *таковы, что* $\det A \neq 0,\ \det(A+\delta A)\neq 0$ *и для некоторой матричной нормы* $\left\|A^{-1}\delta A\right\| < 1$. *Тогда справедлива оценка относительной погрешности вычисления обратной матрицы*

$$\frac{\left\|A^{-1}-(A+\delta A)^{-1}\right\|}{\|A^{-1}\|} \le \frac{\operatorname{cond}(A)}{1-\|A^{-1}\delta A\|}\frac{\|\delta A\|}{\|A\|}. \tag{7.21}$$

Доказательство. Так как для некоторой матричной нормы $\left\|A^{-1}\delta A\right\| < 1$, то в силу (7.15) справедливо

$$\rho\left(A^{-1}\delta A\right) < 1. \tag{7.22}$$

Рассмотрим $A+\delta A = A\left(E+A^{-1}\delta A\right)$. В силу (7.22) и теоремы 7.9 существует $\left(E+A^{-1}\delta A\right)^{-1}$ и верно представление

$$\left(E+A^{-1}\delta A\right)^{-1} = \sum_{k=0}^{\infty}(-1)^k\left(A^{-1}\delta A\right)^k.$$

Тогда

$$\begin{aligned}
\left\|A^{-1}-(A+\delta A)^{-1}\right\| &= \left\|A^{-1}-\left(E+A^{-1}\delta A\right)^{-1}A^{-1}\right\| = \\
&= \left\|A^{-1}-\sum_{k=0}^{\infty}(-1)^k\left(A^{-1}\delta A\right)^k A^{-1}\right\| = \\
&= \left\|\sum_{k=1}^{\infty}(-1)^{k+1}\left(A^{-1}\delta A\right)^k A^{-1}\right\| \le \\
&\le \sum_{k=1}^{\infty}\left\|A^{-1}\delta A\right\|^k\left\|A^{-1}\right\| \overset{(7.22)}{=\!=\!=} \frac{\left\|A^{-1}\delta A\right\|}{1-\left\|A^{-1}\delta A\right\|}\left\|A^{-1}\right\|.
\end{aligned}$$

откуда с учетом (7.20) следует оценка (7.21). □

Следствие 7.3. *Если в условиях теоремы 7.10 матрица* δA *настолько «мала», что*

$$\delta A < \frac{1}{\|A^{-1}\|},$$

то имеет место оценка

$$\frac{\left\|A^{-1}-(A+\delta A)^{-1}\right\|}{\|A^{-1}\|} \le \frac{\operatorname{cond}(A)}{1-\operatorname{cond}(A)\frac{\|\delta A\|}{\|A\|}}\frac{\|\delta A\|}{\|A\|}. \tag{7.23}$$

При малых значениях нормы $\|\delta A\|$ правая часть в (7.23) эквивалентна выражению $\operatorname{cond}(A)\frac{\|\delta A\|}{\|A\|}$. Таким образом, есть все основания считать, что относительная ошибка в обратной матрице имеет одинаковый порядок малости с относительной ошибкой в начальных данных при условии, что величина $\operatorname{cond}(A)$ не слишком велика.

Имея в виду задачу обращения, при больших значениях $\operatorname{cond}(A)$ говорят о *плохой обусловленности* матрицы A (по отношению к матричной норме $\|\cdot\|$). Когда величина $\operatorname{cond}(A)$ мала (близка к единице), говорят о *хорошей обусловленности* матрицы A (по отношению к матричной норме $\|\cdot\|$). Наконец, при $\operatorname{cond}(A) = 1$ матрицу называют *идеально обусловленной* (по отношению к матричной норме $\|\cdot\|$).

Пример 7.12 (Спектральное число обусловленности).

Найдем спектральное число обусловленности матрицы

$$A = \begin{bmatrix} 1 & -1 \\ -1 & 1+\varepsilon \end{bmatrix}, \ \varepsilon > 0.$$

Для этого вычислим

$$A^{-1} = \frac{1}{\varepsilon}\begin{bmatrix} \varepsilon+1 & 1 \\ 1 & 1 \end{bmatrix}, \quad \|A\|_2^2 = \frac{1}{2}(4+2\varepsilon+\varepsilon^2+(2+\varepsilon)\sqrt{4+\varepsilon^2}),$$

$$\left\|A^{-1}\right\|_2^2 = \frac{1}{2\varepsilon^2}\left(4+2\varepsilon+\varepsilon^2+(2+\varepsilon)\sqrt{4+\varepsilon^2}\right),$$

откуда согласно (7.20) получаем

$$\operatorname{cond}_2(A) = \|A\|_2 \left\|A^{-1}\right\|_2 = \frac{1}{2\varepsilon^2}(4+2\varepsilon+\varepsilon^2+(2+\varepsilon)\sqrt{4+\varepsilon^2}).$$

Значит, $\operatorname{cond}_2(A) \underset{\varepsilon\to 0}{\to} \infty$, т.е. матрица A плохо обусловлена.

Пример 7.13* (Возмущение единичной матрицы). Пусть E — единичная матрица, $\|\cdot\|$ — подчиненная матричная норма и $\|\delta E\| < 1$. Покажем, что $E - \delta E$ невырожденная и верна оценка

$$\left\|(E-\delta E)^{-1}\right\| \le \frac{1}{1-\|\delta E\|}.$$

Возьмем произвольный вектор $x \ne 0$. Поскольку

$$1 - \|\delta E\| > 0, \|x\| = \|(x-\delta Ex)+\delta Ex\| \le \|x-\delta Ex\| + \|\delta Ex\|,$$

то

$$\begin{aligned}\|(E-\delta E)\,x\| &= \|x-\delta Ex\| \ge \|x\| - \|\delta Ex\| \ge \\ &\ge \|x\| - \|\delta E\|\,\|x\| \ge (1-\|\delta E\|)\,\|x\| > 0.\end{aligned}$$

Следовательно, если $x \ne 0$, то $(E-\delta E)x \ne 0$, т.е. матрица $E - \delta E$ невырожденная.

Из тождества

$$(E-\delta E)(E-\delta E)^{-1} = E$$

имеем

$$(E-\delta E)^{-1} = E + \delta E(E-\delta E)^{-1}.$$

Отсюда с учетом (7.12) имеем

$$\begin{aligned}\|(E-\delta E)^{-1}\| &\le \|E\| + \|\delta E\|\left\|(E-\delta E)^{-1}\right\| = \\ &= 1 + \left\|(E-\delta E)^{-1}\right\|\|\delta E\|.\end{aligned}$$

Из этого неравенства следует решение задачи (часто ее называют задачей о возмущении единичной матрицы).

Пример 7.14* (Возмущение обратной матрицы). Пусть E — единичная матрица, $\|\cdot\|$ — подчиненная матричная норма и $\|\delta E\| < 1$. Получим оценку отклонения матрицы E от матрицы $(E-\delta E)^{-1}$.

Из $(E-\delta E)^{-1} = E+\delta E(E-\delta E)^{-1}$ (см. пример 7.13) получим

$$E - (E-\delta E)^{-1} = -\delta E(E-\delta E)^{-1}.$$

Отсюда в силу неравенства из примера 7.13 имеем

$$\left\|E-(E-\delta E)^{-1}\right\| \le \|\delta E\|\left\|(E-\delta E)^{-1}\right\| \le \frac{\|\delta E\|}{1-\|\delta E\|}.$$

Пример 7.15 (Плохо обусловленная матрица).

Для матрицы

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1+\varepsilon \end{bmatrix}, \ \varepsilon > 0,$$

имеем

$$\det A = \varepsilon \neq 0, \quad A^{-1} = \frac{1}{\varepsilon}\begin{bmatrix} 1+\varepsilon & -1 \\ -1 & 1 \end{bmatrix}.$$

$$\|A\|_\infty = 2+\varepsilon, \quad \left\|A^{-1}\right\|_\infty = \frac{1}{\varepsilon}(2+\varepsilon),$$

$$\mathrm{cond}_\infty A = \frac{1}{\varepsilon}(2+\varepsilon)^2 \underset{\varepsilon\to 0}{\to} \infty.$$

Т.о., матрица A плохо обусловлена по отношению к ∞-норме, что подтверждает ее «слабую» обратимость или «почти вырожденность» при малых ε.

Приложение числа обусловленности к оценке точности решения СЛАУ

Используя число обусловленности матрицы, можно получить априорные оценки точности решения системы линейных алгебраических уравнений.

Пусть дана система линейных алгебраических уравнений

$$Ax = b,\ A \in \mathbb{C}^{n\times n},\ b \in \mathbb{C}^n,\ \det A \neq 0. \tag{7.24}$$

Вследствие ошибок округления в результате решения получается приближенное решение $\tilde{x}$, которое можно рассматривать как точное решение возмущенной системы

$$(A+\delta A)\tilde{x} = b + \delta b,\ A, \delta A \in \mathbb{C}^{n\times n},\ B, \delta B \in \mathbb{C}^n. \tag{7.25}$$

Для оценки того, насколько приближенное решение $\tilde{x}$ отличается от точного решения x, вводится понятие меры такого отличия. В качестве меры используются нормы векторов и согласованные нормы матриц.

Определение 7.16. Пусть задана какая-то векторная норма. Говорят, что число $\|x-\tilde{x}\|$ есть *абсолютная ошибка в* $\tilde{x}$. Если $x \neq 0$, то число

$$\frac{\|x-\tilde{x}\|}{\|x\|}$$

называют *относительной ошибкой в* $\tilde{x}$. Вектор

$$r = x - \tilde{x}$$

называют *невязкой* решения.

Теорема 7.11. *Пусть x — точное решение уравнения (7.24), а $\tilde{x}$ — точное решение уравнения (7.25). Если $\left\|A^{-1}\delta A\right\| < 1$, то при $b \neq 0$ для векторной нормы, согласованной с матричной, имеет место оценка*

$$\frac{\|x-\tilde{x}\|}{\|x\|} \leq \frac{\mathrm{cond}(A)}{1-\|A^{-1}\delta A\|}\left(\frac{\|\delta A\|}{\|A\|} + \frac{\|\delta b\|}{\|b\|}\right). \tag{7.26}$$

Если $\left\|A^{-1}\right\|\|\delta A\| < 1$, то имеет место более грубая оценка

$$\frac{\|x-\tilde{x}\|}{\|x\|} \leq \frac{\mathrm{cond}(A)}{1-\|A^{-1}\|\,\|\delta A\|}\left(\frac{\|\delta A\|}{\|A\|} + \frac{\|\delta b\|}{\|b\|}\right). \tag{7.27}$$

Доказательство проводится аналогично доказательству теоремы 7.10 с использованием (7.24),(7.25). □

Упражнение 7.7*. Провести доказательство теоремы 7.11.

Пример 7.16 (Обусловленность СЛАУ).

Пусть дана матрица A порядка n с параметром a. Вычислим $\mathrm{cond}_\infty(A)$. Имеем:

$$A = \begin{bmatrix} 1 & a & 0 & \ldots & 0 & 0 \\ 0 & 1 & a & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 1 & a \\ 0 & 0 & 0 & \ldots & 0 & 1 \end{bmatrix},$$

$$A^{-1} = \begin{bmatrix} 1 & -a & a^2 & \dots & (-a)^{n-2} & (-a)^{n-1} \\ 0 & 1 & -a & \dots & (-a)^{n-3} & (-a)^{n-2} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 & -a \\ 0 & 0 & 0 & \dots & 0 & 1 \end{bmatrix}.$$

Тогда

$$\|A\|_\infty = 1 + |a|,$$

$$\left\|A^{-1}\right\|_\infty = 1 + |a| + |a|^2 + \ldots + |a|^{n-1} = \frac{|a|^n - 1}{|a| - 1}, |a| \neq 1,$$

$$\left\|A^{-1}\right\|_\infty = n, \ \ |a| = 1;$$

$$\operatorname{cond}_\infty(A) = \begin{cases} \frac{(|a|+1)(|a|^n-1)}{|a|-1}, & |a| \neq 1, \\ (|a|+1)\,n, & |a| = 1. \end{cases}$$

Отсюда видно, что матрица A плохо обусловлена при $|a| > 1$ и хорошо обусловлена при $|a| \le 1$. Например, при $n = 20$, $a = 5$ будем иметь $\operatorname{cond}_\infty(A) \approx 10^{14}$.

Оценим возмущение в компоненте x_1 решения системы $Ax = b$, если компонента b_n вектора b возмущена на ε.

Точное вычисленное значение $\tilde{x}_1$ компоненты x_1 имеет вид

$$\tilde{x}_1 = b_1 - ab_2 + \ldots + (-a)^{n-2} b_{n-1} + (-a)^{n-1}(b_n + \varepsilon) = x_1 + (-a)^{n-1}\varepsilon.$$

Следовательно, при $|a| > 1$ возмущение в b_n увеличивается в компоненте x_1 в $|a|^{n-1}$ раз, а при $|a| < 1$ во столько же раз уменьшается.

Задачи и упражнения

1. Доказать, что для любой матричной нормы справедливо
   a) $|\,\|A\| - \|B\|\,| \le \|A - B\|$;
   b) $\left\|A^k\right\| \le \|A\|^k$, $\forall k = 1, 2, \ldots, A \in \mathbb{C}^{n\times n}$.
2. Показать на примере, что неравенство b) из предыдущего упражнения, вообще говоря, неверно для аддитивной матричной нормы.

3. Доказать, что если $\left\|A^{-1}\right\| \|F\| < 1$, то $\|F\| < \|A\|$.
4. Доказать, что если S — невырожденная $n \times n$-матрица и $\|\cdot\|$ — матричная норма, то функция $\|A\|_S \equiv \left\|SAS^{-1}\right\|$ будет матричной нормой.
5. Показать, что нормы 1*, 2*, 5*, 6* удовлетворяют всем аксиомам матричной нормы.
6. Доказать, что для любой $A \in \mathbb{C}^{n\times m}$ справедливо
$$\|A\|_E \le \|A\|_M.$$
7. Показать на примере, что функция $\|A\| \equiv \max\limits_{i,j} |a_{ij}|$ не удовлетворяет кольцевому свойству 4) и, следовательно, не является матричной нормой (сравни с нормой 5*).

8*. Показать, что для матричных норм $\|\cdot\|_1$, $\|\cdot\|_2$, $\|\cdot\|_\infty$ выполнено $\|D\| = \max\limits_{1\le i\le n} |d_i|$ для всех диагональных матриц $D = \operatorname{diag}(d_1, d_2, \ldots, d_n)$.

9. Найти матричные нормы 1* — 6* для матриц
$$A = \tfrac{1}{\sqrt{3}}\begin{bmatrix} 3 & -1 \\ 0 & \sqrt{8} \end{bmatrix},\ B = \begin{bmatrix} i & 0 \\ 1-2i & 1 \end{bmatrix},$$
$$C = \begin{bmatrix} 0 & -2i \\ 1 & 2+i \end{bmatrix},\ D = \begin{bmatrix} 1 & -2 \\ -1 & 2 \end{bmatrix},\ F = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix},$$
$$G = \begin{bmatrix} 4 & -2 & 4 \\ -2 & 1 & -2 \\ 4 & -2 & 4 \end{bmatrix},\ H = \begin{bmatrix} 1 & 2 & 3 \\ 3 & -1 & 2 \\ 5 & 4 & 3 \end{bmatrix},$$
$$K = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}^k, k = 1, 2, \ldots,\ \ L = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$
10. Вычислить нормы 1* — 6* для матрицы A из примера 7.12.

11*. Показать, что для спектральной матричной нормы верно
   a) $\|A\|_2 = \max\limits_{\substack{\|x\|_2=1 \\ \|y\|_2=1}} |y^*Ax|$;
   b) $\|A\|_2 = \|A^*\|_2$;
   c) $\|A^*A\|_2 = \|A^*\|_2^2$;

d) $\left\| \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \right\|_2 = \max\{ \|A\|_2, \|B\|_2 \}$ $(A, B \in \mathbb{R}^{n\times m})$;

e) $\|U^*AV\|_2 = \|A\|_2$, если U, V — унитарные матрицы.

12*. Проверить, что элементы таблицы 7.1 представляют собой наилучшие значения констант c_M, удовлетворяющих неравенству $\|A\|_\alpha \le c_M \|A\|_\beta$ для всех матриц $A \in \mathbb{C}^{n\times n}$.

13. Чему равно $\|E_n\|_E$?

14. Показать, что подчиненную матричную норму (7.10) можно также вычислить при помощи эквивалентных формул

$$\|A\| = \max_{\|x\|=1} \|Ax\| = \max_{\|x\|\le 1} \|Ax\| = \max_{\|x\|\ne 0} \frac{\|Ax\|}{\|x\|} = \max_{\|x\|=1} \frac{\|Ax\|}{\|x\|}.$$

15*. Показать справедливость неравенства

$$\|A\|_2^2 \le \|A\|_1 \|A\|_\infty.$$

16. Для матрицы

$$A = \begin{bmatrix} \frac{1}{2} & 1 \\ 0 & \frac{1}{2} \end{bmatrix}$$

явно вычислить A^k и $\rho(A^k)$ для $k = 2, 3, \ldots$. Показать, что $\rho(A^k) = [\rho(A)]^k$. Что происходит при $k \to \infty$ с элементами матрицы A^k и величинами $\|A^k\|_1$, $\|A^k\|_\infty$, $\|A^k\|_2$?

17. Пусть

$$A = \begin{bmatrix} \frac{1}{2} & 1 \\ -\frac{1}{8} & \frac{1}{2} \end{bmatrix}$$

и последовательность векторов $\{x^{(k)}\} \in \mathbb{C}^2$ задана рекуррентным соотношением $x^{(k+1)} = Ax^{(k)}$, $k = 0, 1, \ldots$. Показать, что $x^{(k)} \to 0$ при $k \to \infty$ независимо от выбора начального приближения $x^{(0)}$.

18. Докажите, что все собственные значения матрицы A расположены в круге комплексной плоскости с центром в нуле и радиусом $\|A\|$.

19. Показать, что для матрицы, обратной к $A \in \mathbb{C}^{n\times n}$, справедливо равенство $\mathrm{cond}(A) = \mathrm{cond}(A^{-1})$.

20. a) Обосновать неравенства $\mathrm{cond}(AB) \le \mathrm{cond}(A)\mathrm{cond}(B)$.

b) Является ли функция $\mathrm{cond}(\cdot)$ матричной или векторной нормой на $\mathbb{C}^{n\times n}$?

21. Пусть $\mathrm{cond}_\alpha(A)$ и $\mathrm{cond}_\beta(A)$– числа обусловленности матрицы A по отношению к нормам $\|\cdot\|_\alpha$ и $\|\cdot\|_\beta$, соответственно. Показать, что существуют положительные константы C_m и C_M, связывающие эти числа неравенствами

$$C_m \mathrm{cond}_\alpha(A) \le \mathrm{cond}_\beta(A) \le C_M \mathrm{cond}_\alpha(A)$$

для всех матриц A, т.е. различные нормы приводят к эквивалентным определениям числа обусловленности.

22*. Вывести неравенство

$$\mathrm{cond}(A) \ge \frac{\max\limits_i |\lambda_i(A)|}{\min\limits_i |\lambda_i(A)|}.$$

23. Если величина определителя A мала (или велика), должна ли матрица A быть плохо обусловлена?

24*. Получить из оценки (7.26) оценку в случае, когда возмущена a) только матрица A; b) только вектор b.

25*. Доказать, что в окрестности любой невырожденной матрицы обратная матрица и решение системы линейных алгебраических уравнений с заданной матрицей являются непрерывными функциями входных данных.

26. Вычислить для заданных матриц число обусловленности по отношению к нормам $\|\cdot\|_1$, $\|\cdot\|_\infty$, $\|\cdot\|_E$, $\|\cdot\|_M$, $\|\cdot\|_{l_1}$, $\|\cdot\|_2$:

$$A = \begin{bmatrix} 0 & -2i \\ 1 & 2+i \end{bmatrix}, \quad B = \begin{bmatrix} -i & 0 \\ 2+i & 1 \end{bmatrix}.$$

27*. Доказать, что евклидова матричная норма не является подчиненной ни к какой векторной норме.

28. Доказать, что для матрицы $A = (a_{ij}) \in \mathbb{R}^{n\times n}$ верно неравенство $|\det A| \le n^2\rho$, где $\rho = \max\limits_{i,j} a_{ij}$.

29. Показать, что если записать матрицу $A \in \mathbb{C}^{n\times n}$ через ее вектор-столбцы $a_i \in \mathbb{C}^n$:

$$A = [a_1, a_2, \dots, a_n] \in \mathbb{C}^{n\times n},$$

то евклидову норму матрицы A можно представить в виде

$$\|A\|_E^2 = \|a_1\|_2^2 + \|a_2\|_2^2 + \ldots + \|a_n\|_2^2 .$$

30. Доказать, что для всякой унитарной матрицы $\|U\|_2 = 1$.

31. Матричная норма $\|\cdot\|$ называется *унитарно инвариантной,* если для любого $A \in \mathbb{C}^{n\times n}$ и любых унитарных матриц $U, V \in \mathbb{C}^{n\times n}$ справедливо $\|A\| = \|UAV\|$. Доказать, что для любых унитарных матриц U, V и любой унитарно инвариантной нормы $\|\cdot\|$ справедливо $\mathrm{cond}_{\|\cdot\|}A = \mathrm{cond}_{\|\cdot\|}UA = \mathrm{cond}_{\|\cdot\|}AV = \mathrm{cond}_{\|\cdot\|}UAV$.

32*. Показать, что число обусловленности невырожденной нормальной матрицы по отношению к спектральной норме задается формулой

$$\mathrm{cond}_{\|\cdot\|_2}A = \rho(A)\rho(A^{-1}) = \left|\frac{\lambda_{\max}(A)}{\lambda_{\min}(A)}\right|.$$

Здесь под максимальным и минимальным собственными значениями понимаются максимальное и минимальное по абсолютной величине.

33. Найти спектральное число обусловленности унитарной матрицы. Что можно сказать об обусловленности унитарной матрицы по отношению к спектральной норме?

34. Какие из следующих соотношений верны для произвольной унитарной матрицы $U \in \mathbb{C}^{n\times n}$, произвольного вектора $x \in \mathbb{C}^n$ и произвольной матрицы $A \in \mathbb{C}^{n\times n}$:

a) $\mathrm{cond}_2(UA) = \mathrm{cond}_2 A$; b) $\mathrm{cond}_\infty(UA) = \mathrm{cond}_\infty A$;

c) $\mathrm{cond}_E(UA) = \mathrm{cond}_E A$; d) $\|x\|_2 = \|Ux\|_2$;

e) $\mathrm{cond}_2(AU) = \mathrm{cond}_2 A$?

Вопросы для повторения

1. Что такое матричная норма?
2. Какую матричную норму называют мультипликативной?
3. Объясните, почему $\|A\|_E = \|A^*\|_E$.
4. Как определяется матричная норма, согласованная с векторной нормой?
5. Докажите, что подчиненная матричная норма согласована с соответствующей векторной нормой.
6. Какие матричные нормы подчинены векторным нормам $\|\cdot\|_1$, $\|\cdot\|_\infty$, $\|\cdot\|_2$?
7. Как соотносятся спектральный радиус и норма для заданной матрицы A?
8. Какие условия сходимости матрицы вам известны?
9. Когда говорят, что матрица хорошо обусловлена (плохо обусловлена, идеально обусловлена) по отношению к заданной матричной норме?
10. Зависит ли обусловленность матрицы от величины ее определителя?
11. Что такое приближенное решение системы линейных алгебраических уравнений? Как его можно трактовать?
12. Что такое абсолютная, относительная ошибка решения системы уравнений?
13. Как можно оценить точность решения системы линейных алгебраических уравнений?
14. Как вычислить вектор невязки для оценки точности решения системы линейных алгебраических уравнений?

Необходимо усвоить

Основные понятия и теоремы

- нормированный вектор
- единичный шар
- матрица
  - сходящаяся
  - идеально, хорошо, плохо обусловленная
- норма
  - векторная
    - – – евклидова, абсолютная, максимальная, Гельдера
    - – – полиэдральная, октаэдрическая, кубическая
  - матричная
    - – – столбцовая, строчная, евклидова, спектральная, M-норма, l_1-норма
    - – – * аддитивная, мультипликативная, обобщенная
    - – – подчиненная (индуцированная) векторной норме
    - – – согласованная с векторной нормой
    - – – * минимальная
    - – – * унитарно инвариантная
- нормы эквивалентные
- * полунорма векторная
- сингулярное число матрицы
- спектр матрицы
- спектральный радиус
- число обусловленности
- * последовательность векторов, сходящаяся по норме
- * теорема о суперпозиции векторных норм

- теорема об эквивалентности векторных норм в конечномерных пространствах
- теорема о непрерывности нормы
- критерий сходимости матрицы
- теорема о представлении евклидовой нормы
- теорема о подчиненной норме

Основные умения и навыки

- вычислять наиболее употребительные векторные и матричные нормы в конечномерных пространствах
- знать, какие матричные нормы подчинены наиболее употребительным векторным нормам
- вычислять число обусловленности заданной матрицы по заданной норме
- * строить единичный шар для наиболее употребительных матричных норм в пространстве $\mathbb{R}^{2\times 2}$.
- находить сингулярные числа заданной матрицы
- оценивать величину спектрального радиуса заданной матрицы с использованием матричных норм.
- * конструировать новые матричные нормы исходя из заданных, применяя суперпозицию норм или невырожденное линейное преобразование векторного пространства
- * исследовать сходимость векторной последовательности по заданной норме
- * находить константы эквивалентности для векторных и матричных норм в конечномерных пространствах
- * находить матричные нормы, подчиненные наиболее употребительным векторным нормам.

Глава 8

Локализация и возмущение собственных значений

Глава посвящена нахождению условий и оценок, позволяющих по элементам матрицы указать ограниченные области комплексной плоскости, содержащие все собственные значения матриц. Это бывает полезно, когда нет необходимости знать точные величины собственных значений матрицы, а достаточно знать принадлежность их некоторым областям, либо в случае, когда предварительная локализация собственных значений упрощает дальнейший анлиз. Локализация собственных значений имеет приложения, например в статистике (критерии регулярности цепи Маркова), численном анализе (оценка скорости сходимости процессов), теории дифференциальных уравнений (устойчивость обыкновенных дифференциальных уравнений, анализ и синтез динамических систем).

Рекомендуемая литература:

[3], [4], [5], [9], [10], [13], [14], [16], [17], [19]

Цель изучения

Изучить простейшие способы локализации спектра матрицы (без его вычисления) и применять области локализации для выяснения свойств матрицы; изучить методы оценки собственных значений матрицы, полученной из матрицы с известными собственными значениями в результате малых возмущений ее элементов.

Следует повторить

- собственные значения, собственные векторы, спектр
- комплексные числа, комплексная плоскость
- модуль комплексного числа, геометрическая интерпретация
- формы записи комплексных чисел, связь между ними
- преобразование подобия и свойства подобных матриц
- теорема Шура
- блочные, треугольные и диагональные матрицы, их собственные значения, операции с ними (умножение на диагональную (блочно-диагональную) матрицу слева и справа)
- обратная к диагональной матрице
- связь спектрального радиуса и нормы матрицы
- определение и критерий сходимости матрицы
- теорема Рэлея—Ритца

8.1. Локализация собственных значений

8.1.1. Теорема Гершгорина и ее следствия

Для произвольной матрицы $A \in \mathbb{C}^{n\times n}$ обозначим

$$R_i\,(A) \equiv \sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}|,\ 1 \le i \le n,$$

— строчные почти-нормы, а также

$$C_j\,(A) \equiv \sum_{\substack{i=1\\ i\neq j}}^{n} |a_{ij}|,\ 1 \le j \le n,$$

— столбцовые почти-нормы матрицы A.

Наиболее полезная и легко применяемая теорема, дающая оценки для собственных значений — теорема Гершгорина.

Теорема 8.1 (Гершгорина строчная). *Все собственные значения матрицы A заключены в объединении n кругов*

$$\bigcup_{i=1}^{n} \Big\{ z \in \mathbb{C} \colon |z - a_{ii}| \le R_i(A) \Big\} \equiv G(A). \tag{8.1}$$

Кроме того, если объединение k из этих кругов есть связная область, не пересекающаяся с остальными $n-k$ кругами, то в ней находится ровно k собственных значений матрицы A.

Доказательство. Докажем первую часть теоремы. Пусть λ — собственное значение матрицы A, и пусть $x = [x_i] \neq 0$ — соответствующий ненулевой собственный вектор: $Ax = \lambda x$. Пусть x_p — компонента вектора x с наибольшей абсолютной величиной такой, что

$$|x_p| \ge |x_i|,\ i = \overline{1,n}, x_p \neq 0.$$

Из $Ax = \lambda x$ следует:

$$\lambda x_p = (\lambda x)_p = (Ax)_p = \sum_{j=1}^{n} a_{pj} x_j = \sum_{\substack{j=1\\ j\neq p}}^{n} a_{pj} x_j + a_{pp} x_p.$$

Следовательно,

$$(\lambda - a_{pp})\, x_p = \sum_{\substack{j=1\\ j\neq p}}^{n} a_{pj} x_j.$$

Воспользуемся неравенством треугольника:

$$|x_p| \cdot |\lambda - a_{pp}| = \Big| \sum_{\substack{j=1\\ j\neq p}}^{n} a_{pj} x_j \Big| \le$$

$$\le \sum_{\substack{j=1\\ j\neq p}}^{n} |a_{pj} x_j| \le |x_p| \sum_{\substack{j=1\\ j\neq p}}^{n} |a_{pj}| = |x_p| R_p\,(A)\,.$$

Так как $|x_p| > 0$, то из полученного неравенства имеем, что для некоторого p верно $|\lambda - a_{pp}| \le R_p$, т.е. λ находится в замкнутом круге с центром a_{pp} и радиусом $R_p\,(A)$. Так как мы не знаем, какое именно p соответствует данному λ (разве что известен соответствующий собственный вектор, но тогда мы точно знали бы λ и не нуждались в его локализации), то можно лишь утверждать, что λ принадлежит объединению всех таких кругов, т.е. области (8.1).

Докажем вторую часть теоремы. Представим матрицу A в виде

$$A = B + D,\ D = \mathrm{diag}\big\{a_{11}, a_{22}, \ldots a_{nn}\big\},$$

$$B = \begin{bmatrix} 0 & a_{12} & \ldots & a_{1n} \\ a_{21} & 0 & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & 0 \end{bmatrix}.$$

Наряду с A рассмотрим матрицу $A_\varepsilon = D + \varepsilon B, \quad \varepsilon \ge 0:$

$$A_\varepsilon = \begin{bmatrix} a_{11} & \varepsilon a_{12} & \ldots & \varepsilon a_{1n} \\ \varepsilon a_{21} & a_{22} & \ldots & \varepsilon a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ \varepsilon a_{n1} & \varepsilon a_{n2} & \ldots & a_{nn} \end{bmatrix}.$$

Значит, $A_0 = D, A_1 = A$. Строчные почти-нормы матриц A и A_ε связаны следующим образом:

$$R_i\,(A_\varepsilon) = R_i\,(\varepsilon B) = \varepsilon R_i\,(B) = \varepsilon R_i\,(A)\,.$$

Пусть первые k кругов составляют связную область G_k , не пересекающуюся с дополнительной областью G_k^c, образованной прочими $n-k$ кругами (рис. 8.1).

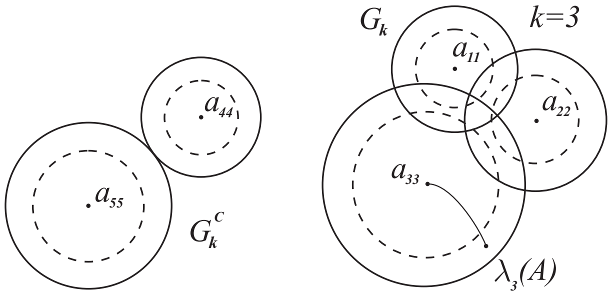


Рис. 8.1

Объединение первых k кругов матрицы A_ε при всех $\varepsilon \in [0,1]$ содержится в связном множестве $G_k \equiv G_k(1)$. Действительно,

$$G_k(\varepsilon) = \bigcup_{i=1}^{n} \Big\{ z \in \mathbb{C} : |z - a_{ii}| \le R_i(A_\varepsilon) = \varepsilon R_i(A_\varepsilon) \Big\} \subseteq G_k,$$

но само $G_k(\varepsilon)$ не обязано быть связным для всех ε . Кроме того, никакая из дополнительных областей $G_k^c(\varepsilon) = G_n(\varepsilon) \backslash G_k(\varepsilon)$ не может пересекаться с G_k.

Покажем, что в области G_k имеется не менее k собственных значений матрицы A. Рассмотрим собственные значения

$$\lambda_i(A_0) = a_{ii}, \quad \lambda_i(A_\varepsilon) = \lambda_i(A_0 + \varepsilon B), \quad \varepsilon > 0,$$
$$\lambda_i(A) = \lambda_i(A_1), \quad i = \overline{1,k}.$$

Поскольку собственные значения непрерывно зависят от элементов матрицы A, то при $\varepsilon \in [0,1]$ функция

$$y(\varepsilon) = \lambda_i\left(A_0 + \varepsilon B\right) = \lambda_i(A_\varepsilon),\ 0 \le \varepsilon \le 1,$$

задаёт на комплексной плоскости непрерывную прямую, соединяющую точки $\lambda_i(A_0) = a_{ii}$ (центр круга) и $\lambda_i(A_1) = \lambda_i(A)$ (собственное значение матрицы A).

В силу включения

$$\lambda_i(A_\varepsilon) \in G_k(\varepsilon) \subseteq G_k, \quad \varepsilon \in [0,1]$$

эти кривые (в том числе и их концы $\lambda_i(A)$) принадлежат области G_k. Таким образом, при каждом $\varepsilon \in [0,1]$ в $G_k(\varepsilon)$ содержится по крайней мере k собственных значений матрицы $A(\varepsilon)$.

Покажем, что больше чем k собственных значений быть не может. Действительно, остальные $n-k$ собственных значений матрицы $A(\varepsilon)$ связаны непрерывными линиями с точками $\lambda_i(A_0)$, которые расположены вне связной области G_k и лежат в дополнительной области G_k^c. Из-за непрерывности они не могут преодолеть разрыв между G_k^c и G_k. □

Область $G(A)$ (8.1) называют *(строчной) областью Гершгорина,* отдельные круги в $G(A)$ — *кругами Гершгорина,* а их границы – *окружностями Гершгорина.*

Если какой-либо круг Гершгорина изолирован от остальных, то он содержит точно одно собственное значение.

Пример 8.1 (Строчная область Гершгорина). Изобразим область Гершгорина для матрицы

$$A = \begin{bmatrix} 0 & 2 \\ -3 & 2i \end{bmatrix}.$$

Она образуется двумя кругами: кругом $|z| \le 2$ с центром $z = a_{11} = 0$ и радиусом $R_1 = |2| = 2$ и кругом $|z - 2i| \le 3$ с центром $z = a_{22} = 2i$ и радиусом $R_2 = |-3| = 3$ (см. рис. 8.2).

i	***Круг*** i	***Центр*** z_i	***Радиус*** R_i
1	$\|z\| \le 2$	0	2
2	$\|z-2i\| \le 3$	$2i$	3

Матрица A имеет собственные значения $\lambda_1 = i(1+\sqrt{7}) \approx 3,65i$ и $\lambda_2 = i(1-\sqrt{7}) \approx -1,65i$, которые лежат в области Гершгорина и отмечены на рисунке 8.2 кружочками.

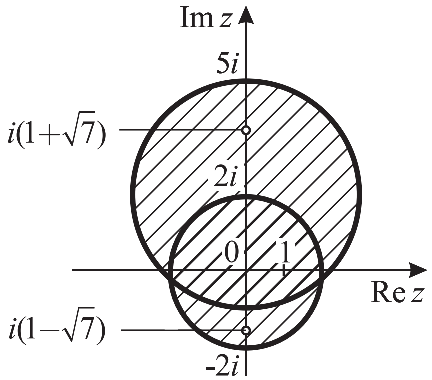


Рис. 8.2

Пример 8.2 (Строчная область Гершгорина). Для матрицы

$$B = \begin{bmatrix} 0 & 1 & 0 \\ 2 & 1 & 0 \\ 1 & 0 & -3 \end{bmatrix}$$

строчная область Гершгорина состоит из трех кругов (рис. 8.3):

i	***Круг*** i	***Центр*** r_i	***Радиус*** R_i
1	$\lvert z\rvert \le 1$	0	$\lvert 1\rvert + \lvert 0\rvert = 1$
2	$\lvert z-1\rvert \le 2$	1	$\lvert 2\rvert + \lvert 0\rvert = 2$
3	$\lvert z+3\rvert \le 1$	-3	$\lvert 1\rvert + \lvert 0\rvert = 1$

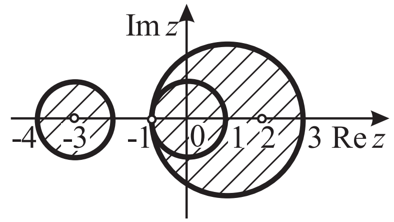


Рис. 8.3

Причем третий круг изолирован от первых двух, и, значит, содержит в точности одно собственное значение ($\lambda_1 = -3$). Два других собственных значения ($\lambda_2 = -1$ и $\lambda_3 = 2$) лежат в объединении первых двух кругов.

Поскольку A и A^T имеют одни и те же собственные значения, то применение к матрице A^T теоремы Гершгорина дает следующий результат.

Теорема 8.2 (Гершгорина (столбцовая)). *Все собственные значения матрицы $A \in \mathbb{C}^{n\times n}$ принадлежат объединению n кругов (столбцовой области Гершгорина)*

$$\bigcup_{j=1}^{n} \{ z \in \mathbb{C} : |z - a_{jj}| \le C_j(A) \} \equiv G(A^T). \tag{8.2}$$

Кроме того, если k из этих кругов образуют область, изолированную от остальных $n-k$ кругов, то в ней находится ровно k собственных значений матрицы A.

Пример 8.3 (Столбцовая область Гершгорина).

Столбцовая область Гершгорина для матрицы $A \in \mathbb{C}^{2\times 2}$ из примера 8.1 изображена на рис. 8.4 и состоит из двух кругов:

j	***Круг*** j	***Центр*** z_j	***Радиус*** C_j
1	$\lvert z\rvert \le 3$	0	$\lvert -3\rvert = 3$
2	$\lvert z-2i\rvert \le 2$	$2i$	$\lvert 2\rvert = 2$

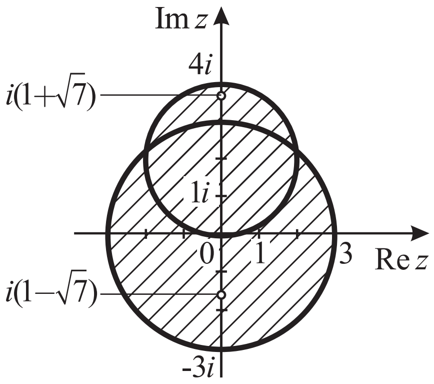


Рис. 8.4

Пример 8.4 (Столбцовая область Гершгорина).

Столбцовая область Гершгорина для матрицы $B \in \mathbb{C}^{3\times 3}$ из примера 8.2 изображена на рис. 8.5 и состоит из трех кругов (третий круг вырождается в точку):

j	***Круг*** j	***Центр*** r_j	***Радиус*** C_j
1	$\lvert z\rvert \le 3$	0	$\lvert 0\rvert + \lvert 0\rvert = 0$
2	$\lvert z-1\rvert \le 1$	1	$\lvert 1\rvert + \lvert 0\rvert = 1$
3	$\lvert z+3\rvert \le 0$	-3	$\lvert 0\rvert + \lvert 0\rvert = 0$

Поскольку все собственные значения матрицы $A \in \mathbb{C}^{n\times n}$ принадлежат одновременно $G(A)$ и $G(A^T)$, то пересечение строчной и столбцовой областей Гершгорина $G(A)\bigcap G(A^T)$ содержит все собственные значения матрицы A.

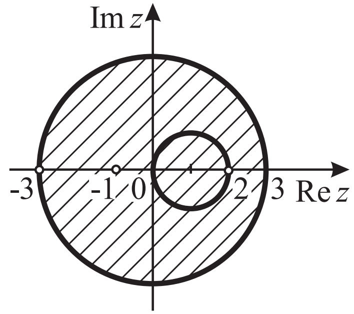


Рис. 8.5

Пример 8.5 (Область локализации Гершгорина).

Пересечение строчной и столбцовой областей Гершгорина для матриц из примеров 8.1 и 8.2 изображено на рис. 8.6.

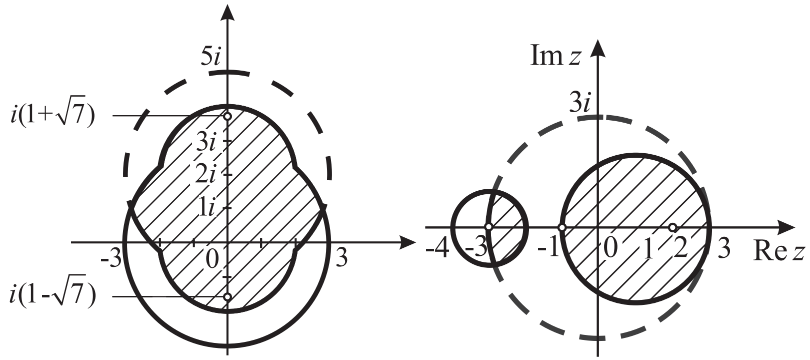


Рис. 8.6

Пусть у матрицы A элементы главной диагонали и коэффициенты характеристического многочлена — вещественные. Если круги Гершгорина попарно не пересекаются, то все собственные значения матрицы A — вещественные.

Упражнение 8.1. Доказать предыдущее утверждение.

Пример 8.6 (Локализация вещественного спектра).

Для матрицы

$$A = \begin{bmatrix} 2 & i \\ 2i & -2 \end{bmatrix}$$

с вещественными диагональными элементами характеристический многочлен $p_A(\lambda) = \lambda^2 - 2$ имеет вещественные коэффициенты, круги Гершгорина $|z-2| \le 1$ и $|z+2| \le 2$ не пересекаются (см. рис. 8.7).

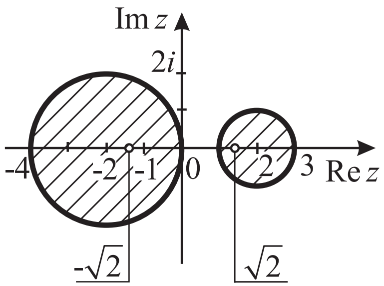


Рис. 8.7

Следовательно, все собственные значения матрицы A — вещественные. Непосредственное вычисление дает

$$\lambda_1 = \sqrt{2} \approx 1,4; \ \lambda_2 = -\sqrt{2} \approx -1,4.$$

8.1.2. Уточнение оценок собственных значений

Если относительно матрицы имеется некоторая дополнительная информация, в силу которой собственные значения принадлежат (или не принадлежат) конкретным множествам, то в сочетании с теоремой Гершгорина эта информация может привести к более точной локализации собственных значений.

Определяющие свойства некоторых типов матриц и свойства их собственных значений приведены в таблице

Название матрицы	***Определяющее условие***	***Собственные значения***
эрмитова	$A = A^*$	вещественные ($\operatorname{Im}\lambda_i = 0$)
косоэрмитова	$A = -A^*$	чисто мнимые ($\operatorname{Re}\lambda_i = 0$)
унитарная	$A = A^* = A^*A = E_n$	по модулю равны 1 ($\lvert\lambda_i\rvert = 1$)
нильпотентная	$\exists p \ge 2:$ $A^{p-1} \ne 0, \ A^p = 0$	равны 0
идемпотентная	$A^2 = A$	равны 0 или 1
стохастическая	$A = (a_{ij}) \ge 0,$ $\sum\limits_{j=1}^{n} a_{ij} = 1$	лежат в единичном круге ($\lvert\lambda_i\rvert \le 1$) и по крайней мере одно из них равно 1

Пример 8.7 (Локализация спектра эрмитовой матрицы).

Матрица

$$A = \begin{bmatrix} 1 & -2i \\ 2i & 0 \end{bmatrix}$$

— эрмитова, поэтому все ее собственные значения лежат на вещественной оси и из области Гершгорина (рис. 8.8) вырезается вещественный отрезок $[-2, 3]$, содержащий оба собственных значения $\lambda_1 \approx 1,56$; $\lambda_2 \approx 2,56$.

Пример 8.8 (Изолированный круг Гершгорина).

Для матрицы

$$A = \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 12 & 0 & -4 \\ -1 & 0 & -1 & 0 \\ 0 & 5 & 0 & 0 \end{bmatrix} \in \mathbb{R}^{4\times 4}$$

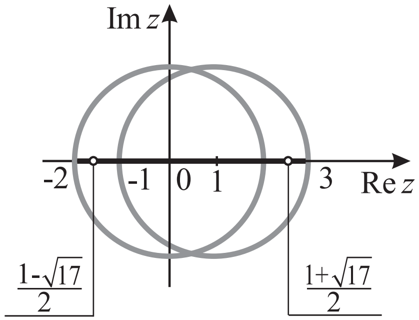


Рис. 8.8

строчная область Гершгорина изображена на рис. 8.9.

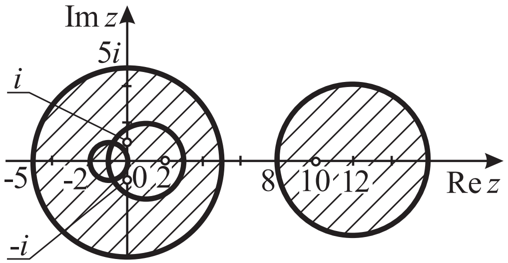


Рис. 8.9

Так как один круг изолирован от остальных трех, то в нем содержится единственное собственное значение, а остальные три лежат в объединении трех оставшихся кругов. Поскольку собственные значения вещественных матриц образуют пары комплексно-сопряженных чисел, то собственное значение в изолированом круге должно быть вещественным и должно существовать по крайней мере еще одно вещественное собственное значение в объединении трех других кругов. Вычисление собственных значений матрицы A дает $\sigma(A) = \{\pm i,\ 2,\ 10\}$.

Поскольку спектральный радиус не превосходит модуля самой удаленной от начала координат точки пересечения строч-

ной и столбцовой областей Гершгорина, то для матрицы $A \in \mathbb{C}^{n\times n}$ справедлива оценка спектрального радиуса

$$\rho(A) \le \min\{\max_i \sum_{j=1}^{n} |a_{ij}|,\ \max_j \sum_{i=1}^{n} |a_{ij}|\}. \tag{8.3}$$

Упражнение 8.2. Дать полное доказательство оценки 8.3.

Пример 8.9 (Оценка спектрального радиуса).

Для матрицы A из примера 8.1 имеем оценку (8.3)

$$\rho(A) \le \min\Big\{\max\big\{|0| + |2|; |-3| + |2i|\big\},$$

$$\max\big\{|0| + |-3|; |2| + |2i|\big\}\Big\} =$$

$$= \min\{\max\{2; 5\} \max\{3; 4\}\} = \min\{5, 4\} = 4.$$

Непосредственное вычисление дает

$$\rho(A) = \max\{\ |i(1-\sqrt{7})|, |i(1+\sqrt{7})|\} = |i(1+\sqrt{7})| \approx 3,65.$$

Для уточнения оценок собственных значений можно использовать преобразование подобия.

Для любой обратимой матрицы S матрица $S^{-1}AS$ имеет те же собственные значения, что и A: $\sigma\left(S^{-1}AS\right) = \sigma(A)$. Поэтому теорему Гершгорина можно применить к $S^{-1}AS$; возможно, что при подходящем выборе S удастся получить более точные оценки для собственных значений матрицы A. Удобно выбирать

$$S = D = \mathrm{diag}(p_1, p_2, \ldots, p_n),$$

где все $p_i > 0$. Легко вычислить, что

$$D^{-1}AD = \left[\frac{p_j a_{ij}}{p_i}\right].$$

Теорема Гершгорина, записанная для $D^{-1}AD$ и для транспонированной к ней матрице, дает следующий результат.

Следствие 8.1. *Пусть $A \in \mathbb{C}^{n\times n}$, и пусть $p_1, p_2, \ldots, p_n$ — произвольные положительные числа. Все собственные значения матрицы A принадлежат каждой из двух областей*

$$\bigcup_{i=1}^{n}\left\{z \in \mathbb{C} : |z - a_{ii}| \le \frac{1}{p_i}\sum_{\substack{j=1\\ j\ne i}}^{n} p_j\,|a_{ij}|\right\} = G\left(D^{-1}AD\right), \qquad (8.4)$$

$$\bigcup_{j=1}^{n}\left\{z \in \mathbb{C} : |z - a_{jj}| \le p_j\sum_{\substack{j=1\\ j\ne i}}^{n} \frac{1}{p_i}\,|a_{ij}|\right\} = G\left[\left(D^{-1}AD\right)^T\right]. \qquad (8.5)$$

Пример 8.10 (Уточнение оценок спектра). Матрица

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}$$

имеет собственные значения 1 и 2. Прямое применение теоремы Гершгорина дает грубые оценки собственных значений (см. рис. 8.10(a)). В то же время гибкость, обеспечиваемая дополнительными параметрами $p_1, p_2, \ldots, p_n$ из (8.4),(8.5), позволяет получить сколь угодно точные оценки (рис. 8.10(b)).

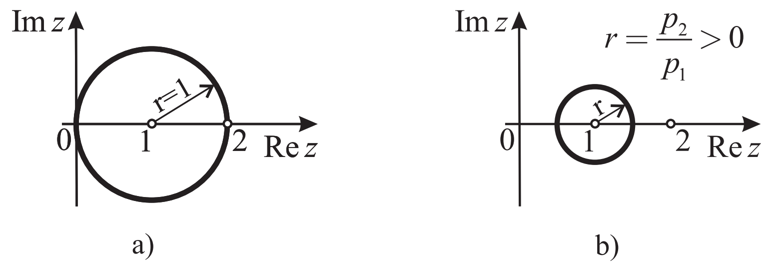


Рис. 8.10

Приведем еще некоторые оценки собственных значений матрицы A.

8.1.3*. Другие области локализации

Теорема 8.3 (Неравенство Шура). *Если $A = (a_{ij}) \in \mathbb{C}^{n\times n}$ имеет собственные значения $\lambda_1, \ldots, \lambda_n$, то*

$$\sum_{i=1}^{n} |\lambda_i|^2 \le \sum_{i,j=1}^{n} |a_{ij}|^2, \qquad (8.6)$$

причем равенство имеет место тогда и только тогда, когда матрица A нормальная.

Доказательство. Согласно теореме Шура (стр. 146) существует унитарная матрица $U \in \mathbb{C}^{n\times n}$, такая, что

$$U^*AU = T,$$

где T — верхняя треугольная матрица с диагональными элементами $t_{ii} = \lambda_i$, $i = \overline{1,n}$. Тогда

$$T^* = (U^*AU)^* = U^*A^*U, \quad t^*_{ii} = \overline{\lambda}_i,$$

а значит

$$TT^* = U^*AUU^*A^*U = U^*AA^*U,$$

т.е. матрицы AA^* и TT^* подобны.

Поскольку след у подобных матриц равен, то

$$\operatorname{tr} AA^* = \operatorname{tr} TT^*. \qquad (8.7)$$

Далее с учетом (2.9) и (8.7) справедливы соотношения

$$\sum_{i,j=1}^{n} |a_{ij}|^2 = \operatorname{tr} AA^* = \operatorname{tr} TT^* = \sum_{i,j=1}^{n} |t_{ij}|^2 =$$

$$= \sum_{i=1}^{n} |\lambda_i|^2 + \sum_{i<j}^{n} |t_{ij}|^2 \ge \sum_{i=1}^{n} |\lambda_i|^2,$$

что и доказывает (8.6). Равенство возможно тогда и только тогда, когда $\sum\limits_{i<j}^{n} |t_{ij}|^2 = 0$, т.е. тогда и только тогда, когда матрица A — нормальная. □

Для матрицы $A \in \mathbb{C}^{n\times n}$ положим

$$B = (b_{ij}) = \frac{(A + A^*)}{2}\ , \quad C = (c_{ij}) = \frac{(A - A^*)}{2i}.$$

Очевидно, что обе матрицы B и C являются эрмитовыми. Будем обозначать собственные значения матриц A, B и C соответственно

$$\lambda_1, \dots, \lambda_n \ (|\lambda_1| \le \dots \le |\lambda_n|),$$

$$\mu_1 \le \dots \le \mu_n \ \text{ и } \ \nu_1 \le \dots \le \nu_n.$$

Теорема 8.4 (Бендиксона). *Если $A \in \mathbb{C}^{n\times n}$, то*

$$\mu_1 \le \operatorname{Re}\lambda_k \le \mu_n,$$
$$\nu_1 \le \operatorname{Im}\lambda_k \le \nu_n.$$

Доказательство. Пусть x — нормированный собственный вектор матрицы A, соответствующий собственному значению λ_k :

$$Ax = \lambda_k x, \quad x^*x = 1. \tag{8.8}$$

Тогда

$$(Ax)^*x = x^*A^*x = (\lambda_k x)^*x = \overline{\lambda}_k x^*x. \tag{8.9}$$

Так как матрица B эрмитова, то для нее справедлива теорема Рэлея—Ритца (стр. 173):

$$x^*Bx = \frac{1}{2}x^*(A + A^*)x = \frac{1}{2}(x^*Ax + x^*A^*x) \overset{(8.8),(8.9)}{=\!=\!=\!=\!=}$$
$$= \frac{1}{2}(\lambda_k + \overline{\lambda}_k)x^*x = \frac{1}{2}\cdot 2\operatorname{Re}\lambda_k = \operatorname{Re}\lambda_k.$$

Отсюда с учетом первого соотношения (6.6) теоремы Рэлея—Ритца имеем

$$\mu_1 x^*x \le x^*Bx \le \mu_n x^*x \Leftrightarrow \mu_1 \le \operatorname{Re}\lambda_k \le \mu_n.$$

Аналогично, применяя теорему Рэлея—Ритца к эрмитовой матрице C, получим второе соотношение теоремы. □

Теорема 8.5 (Гирша). *Если $A \in \mathbb{C}^{n\times n}$, то*

$$|\,\lambda_i| \le n \cdot \max_{i,j} |\,a_{ij}\,| \equiv \|\,A\,\|_M\,, \tag{8.10}$$

$$|\operatorname{Re}\lambda_i| \le n \cdot \max_{i,j} |\,b_{ij}\,| \equiv \|\,B\,\|_M\,, \tag{8.11}$$

$$|\operatorname{Im}\lambda_i| \le n \cdot \max_{i,j} |\,c_{ij}\,| \equiv \|\,C\,\|_M\,. \tag{8.12}$$

Доказательство. Согласно теореме Шура существуют матрицы U — унитарная, T — треугольная, такие, что

$$U^*AU = T, \quad t_{ii} = \lambda_i.$$

Тогда верно:

$$U^*A^*U = T^*,$$
$$U^*\left(\frac{A + A^*}{2}\right)U = \frac{1}{2}(T + T^*) \Rightarrow B = \frac{1}{2}U(T + T^*)U^*,$$
$$U^*\left(\frac{A - A^*}{2i}\right)U = \frac{1}{2i}(T - T^*) \Rightarrow C = \frac{1}{2i}U(T - T^*)U^*. \tag{8.13}$$

Для диагональных элементов матриц верно

$$(T^*)_{ii} = \overline{\lambda}_i,\ (T + T^*)_{ii} = 2\operatorname{Re}\lambda_i,\ (T - T^*)_{ii} = 2\operatorname{Im}\lambda_i. \tag{8.14}$$

Согласно (2.9),(8.13) с учетом инвариантности следа относительно преобразования подобия имеем:

$$\sum_{i,j=1}^{n} |b_{ij}|^2 = \operatorname{tr} BB^* = \operatorname{tr}\left(\frac{T + T^*}{2}\right)\left(\frac{T + T^*}{2}\right)^* =$$
$$= \sum_{i=1}^{n}\left|\frac{t_{ii} + \overline{t}_{ii}}{2}\right|^2 + \sum_{i\ne j}\left|\frac{t_{ij} + \overline{t}_{ij}}{2}\right|^2 \overset{(8.14)}{\ge} \sum_{i=1}^{n}|\operatorname{Re}\lambda_i|^2.$$

Тогда

$$|\operatorname{Re}\lambda_i| \le \sqrt{\sum_{i=1}^{n}|\operatorname{Re}\lambda_i|^2} \le \sqrt{\sum_{i,j=1}^{n}|b_{ij}|^2} \le n\cdot\max|b_{ij}|.$$

Аналогично доказывается (8.12). Неравенство (8.10) следует из (7.15) и определения M-нормы. □

Пример 8.11 (Оценка Гирша). Собственными значениями матрицы

$$A = \begin{bmatrix} 7+3i & -4-6i & -4 \\ -1-6i & 7 & -2-6i \\ 2 & 4-6i & 13-3i \end{bmatrix}$$

являются

$$\lambda_1 = 9,\ \lambda_2 = 9+9i,\ \lambda_3 = 9-9i,$$

(на рис. 8.11 обозначены кружочками),

$$|\lambda_1| = 9,\ |\lambda_2| = |\lambda_3| \approx 12{,}73.$$

Теорема Гирша дает $|\lambda_i| \le 40{,}03$, $|\mathrm{Re}(\lambda_i)| \le 39$, $|\mathrm{Im}(\lambda_i)| \le 20{,}12$. Круги Гершгорина $|z-7-3i| \le 11{,}21$, $|z-7| \le 12{,}4$, $|z-13+3i| \le 9{,}21$ изображены на рис. 8.11.

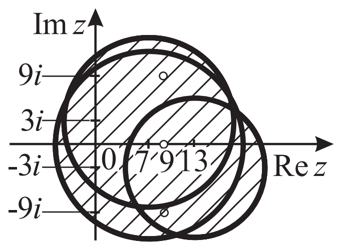


Рис. 8.11

Теорема 8.6 (Островского). *Для любого $\alpha \in [0,\ 1]$ все собственные значения матрицы A принадлежат объединению n кругов*

$$\bigcup_{i=1}^{n} \Big\{ z \in \mathbb{C} : |z - a_{ii}| \le R_i^{\alpha} C_i^{1-\alpha} \Big\}.$$

*Доказательство**. Будем считать $\alpha \in [0,1]$, так как случай $\alpha = 0$, $\alpha = 1$ соответствует теоремам Гершгорина.

Далее будем считать, что $R_i > 0$ для любого i. В противном случае можно было бы возмутить A внесением малых ненулевых элементов в те строки, для которых $R_i = 0$. Для возмущенной матрицы область локализации (8.1) больше, чем область локализации для A (так как центры одинаковые, а радиус больше), и нужный результат получится в пределе при возмущении, стремящемся к нулю.

Пусть λ — собственное значение матрицы A, x — соответствующий ему ненулевой собственный вектор, т.е. $Ax = \lambda x$, $x = [x_i] \ne 0$. Тогда для каждого $i = \overline{1,n}$ с учетом *неравенства Гёльдера* $\sum\limits_i x_i y_i \le (\sum\limits_i x_i^p)^{\frac{1}{p}} (\sum\limits_i y_i^q)^{\frac{1}{q}}$ имеем:

$$\lambda x_i = \sum_{j=1}^{n} a_{ij} x_j \Rightarrow (\lambda - a_{ii}) x_i = \sum_{\substack{j=1 \\ j \ne i}}^{n} a_{ij} x_j \Rightarrow |\lambda - a_{ii}| \cdot |x_i| =$$

$$= |\sum_{\substack{j=1 \\ j \ne i}}^{n} a_{ij} x_j| \le \sum_{\substack{j=1 \\ j \ne i}}^{n} |a_{ij}| \cdot |x_j| = \sum_{\substack{j=1 \\ j \ne i}}^{n} |a_{ij}|^{\alpha} \left\{ |a_{ij}|^{1-\alpha} |x_j| \right\} \le$$

$$\le \Bigg[\sum_{\substack{j=1 \\ j \ne i}}^{n} \left\{ |a_{ij}|^{\alpha} \right\}^{\frac{1}{\alpha}} \Bigg]^{\alpha} \cdot \Bigg[\sum_{\substack{j=1 \\ j \ne i}}^{n} \left\{ |a_{ij}|^{1-\alpha} |x_j| \right\}^{\frac{1}{1-\alpha}} \Bigg]^{1-\alpha} =$$

$$= R_i^{\alpha} \Bigg[\sum_{\substack{j=1 \\ j \ne i}}^{n} \left\{ |a_{ij}| |x_j|^{\frac{1}{1-\alpha}} \right\} \Bigg]^{1-\alpha} \Rightarrow$$

$$\frac{|\lambda - a_{ii}|}{R_i^{\alpha}} |x_i| \le \Bigg[\sum_{\substack{j=1 \\ j \ne i}}^{n} \left\{ |a_{ij}| |x_j|^{\frac{1}{1-\alpha}} \right\} \Bigg]^{1-\alpha} \Rightarrow$$

$$\left[\frac{|\lambda - a_{ii}|}{R_i^{\alpha}} \right]^{\frac{1}{1-\alpha}} |x_i|^{\frac{1}{1-\alpha}} \le \sum_{\substack{j=1 \\ j \ne i}}^{n} |a_{ij}| \cdot |x_j|^{\frac{1}{1-\alpha}}. \tag{8.15}$$

Суммируя (8.15) по i и меняя справа пределы суммирования, получим:

$$\sum_{i=1}^{n}\left[\frac{|\lambda-a_{ii}|}{R_i^{\alpha}}\right]^{\frac{1}{1-\alpha}}|x_i|^{\frac{1}{1-\alpha}} \le \sum_{j=1}^{n}\sum_{\substack{i=1\\ i\neq j}}^{n}|a_{ij}|\cdot|x_j|^{\frac{1}{1-\alpha}} = \sum_{j=1}^{n}C_j|x_j|^{\frac{1}{1-\alpha}}. \tag{8.16}$$

Если для каждого i такого, что $x_i \neq 0$, будет

$$\left[\frac{|\lambda-a_{ii}|}{R_i^{\alpha}}\right]^{\frac{1}{1-\alpha}} > C_i,$$

то (8.16) не может выполняться. Поэтому хотя бы для одного i выполняется

$$\left[\frac{|\lambda-a_{ii}|}{R_i^{\alpha}}\right]^{\frac{1}{1-\alpha}} \le C_i \Rightarrow \frac{|\lambda-a_{ii}|}{R_i^{\alpha}} \le C_i^{1-\alpha} \Rightarrow |\lambda-a_{ii}| \le R_i^{\alpha}C_i^{1-\alpha}.$$

$\square$

Пример 8.12 (Область Островского).

Для матрицы B из примера 8.2 при $\alpha = \frac{1}{2}$ область Островского задается двумя кругами с центрами в 0 и 1 и радиусами $\sqrt{3} \approx 1,73$ и $\sqrt{2} \approx 1,41$, соответственно. Третий круг вырождается в точку и совпадает с собственным значением $\lambda = -3$ (см. рис. 8.12).

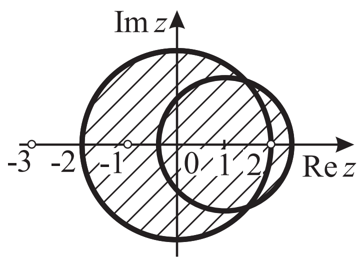


Рис. 8.12

Теорема 8.7 (Брауна). *Если $A \in \mathbb{C}^{n\times n}$ имеет собственные значения $\lambda_1,\ldots,\lambda_n$, и сингулярные числа $\sigma_1 \ge \ldots \ge \sigma_n$, то*

$$\sigma_n \le |\lambda_i| \le \sigma_1. \tag{8.17}$$

Доказательство. Пусть

$$Ax = \lambda_i x, \quad x \in \mathbb{C}, \quad x^*x = \|x\|_2^2 = 1. \tag{8.18}$$

Тогда

$$\|Ax\|_2^2 = \|\lambda_i x\|_2^2 = |\lambda_i|^2\|x\|_2^2 = |\lambda_i|^2.$$

Получаем:

$$|\lambda_i|^2 = \|Ax\|_2^2 = (Ax)^*Ax = x^*A^*Ax.$$

По теореме Рэлея—Ритца для эрмитовой матрицы A^*A верно:

$$\lambda_{\min}(A^*A)x^*x \le x^*(A^*A)x \le \lambda_{\max}(A^*A),$$

откуда с учетом $\lambda_{\min}(A^*A) = \sigma_n^2$, $\lambda_{\max}(A^*A) = \sigma_1^2$ и (8.18) следует справедливость (8.17). $\square$

Теорема 8.8 (Брауэра). *Все собственные значения матрицы A принадлежат объединению $\frac{n(n-1)}{2}$ овалов Кассини*

$$\bigcup_{\substack{i,j=1\\ i\neq j}}^{n}\Big\{ z \in \mathbb{C} : |\, z - a_{ii}\,|\,|\, z - a_{jj}\,| \le R_i\ R_j \Big\}. \tag{8.19}$$

Доказательство.* Пусть λ — собственное значение матрицы A и пусть $Ax = \lambda x$, где $x = [x_i] \neq 0$. Пусть x_p — компонента с наибольшим модулем, так что $|x_p| \ge |x_i|$, $i = \overline{1,n}$, $x_p \neq 0$. Если все остальные компоненты вектора x нулевые, то из условия $Ax = x$ вытекает $a_{pp} = \lambda$. Так как все диагональные элементы матрицы A включены в область (8.19), то всякое собственное значение, отвечающее собственному вектору с единственной ненулевой компонентой, попадает в эту область.

Пусть x имеет по крайней мере две ненулевых компоненты, и x_q — компонента со вторым по величине значением модуля, т.е. $|x_p| \geq |x_q| \geq |x_i|$, $i = \overline{1,n}$, $i \neq p$, $x_p, x_q \neq 0$. Из $Ax = \lambda x$ следует:

$$x_p(\lambda - a_{pp}) = \sum_{\substack{j=1 \\ j\neq p}}^{n} a_{pj}x_j \Rightarrow |x_p| \cdot |\lambda - a_{pp}| = |\sum_{\substack{j=1 \\ j\neq p}}^{n} a_{pj}x_j| \leq$$

$$\leq \sum_{\substack{j=1 \\ j\neq p}}^{n} |a_{pj}||x_j| \leq \sum_{\substack{j=1 \\ j\neq p}}^{n} |a_{pj}||x_q| = R_p|x_q| \Rightarrow$$

$$\Rightarrow |\lambda - a_{pp}| \leq R_p\frac{|x_q|}{|x_p|}. \qquad (8.20)$$

Аналогичным образом получаем

$$|\lambda - a_{qq}| \leq R_q\frac{|x_p|}{|x_q|}. \qquad (8.21)$$

Перемножая (8.20) и (8.21), получим

$$|\lambda - a_{pp}| \cdot |\lambda - a_{qq}| \leq R_p \cdot R_q. \qquad \square$$

8.1.4. Анализ невырожденности матриц

Из всякой теоремы об областях локализации собственных значений можно вывести связанную с ней теорему об обратимости. Нужно лишь, опираясь на результат о локализации, сформулировать условия, которые исключали бы точку 0 из соответствующей области.

Определение 8.1. Матрица $A \in \mathbb{C}^{n\times n}$ называется *матрицей с диагональным преобладанием,* если выполнены *нестрогие условия Адамара:*

$$|a_{ii}| \geq \sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}| = R_i(A), \ i = \overline{1,n}. \qquad (8.22)$$

Если в (8.22) неравенства строгие (*условия Адамара*), то A называется *матрицей со строгим диагональным преобладанием.*

Из теоремы Гершгорина следует, что 0 не является собственным значения, а значит, матрица A обратима, если выполнены условия Адамара. Таким образом, справедлива

Теорема 8.9 (Адамара). *Если A — матрица со строгим диагональным преобладанием, то она обратима.*

Упражнение 8.3. Провести строгое доказательство теоремы 8.9.

Пример 8.13 (Локализация спектра невырожденной эрмитовой матрицы). Рассмотрим эрмитову матрицу

$$A = \begin{bmatrix} 2 & -i \\ i & 3 \end{bmatrix}$$

с положительными диагональными элементами. Так как $|2| > |-i| = 1$, $\quad |3| > |i| = 1$, то A — матрица со строгим диагональным преобладанием. Область Гершгорина матрицы A не содержит точку 0 и лежит в правой комплексной полуплоскости (см. рис. 8.13).

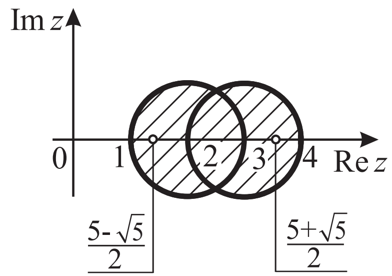


Рис. 8.13

Поэтому матрица A невырожденная, а ее собственные значения должны быть положительны. Непосредственное вычисление дает

$$\lambda_1 = \frac{5+\sqrt{5}}{2} \approx 3,6 > 0, \lambda_2 = \frac{5-\sqrt{5}}{2} \approx 1,4 > 0.$$

В некоторых случаях требование строгого диагонального преобладания как достаточного условия обратимости можно ослабить.

Теорема 8.10. *Если у матрицы $A \in \mathbb{C}^{n\times n}$ все диагональные элементы ненулевые и она является матрицей с диагональным преобладанием, причем для всех, кроме одного, значений $i = 1, 2, \ldots, n$ это свойство выполняется в сильной форме, т.е. $|a_{ii}| > R_i(A)$, то A обратима.*

Доказательство. Из условия теоремы следует, что

$$\exists\, k : |a_{kk}| = R_k, \forall i \neq k : |a_{ii}| > R_i.$$

Положим в (8.4) $p_i = 1,\ i \neq k,\ p_k = 1+\varepsilon, \varepsilon > 0$. Тогда

$$\frac{1}{p_k}\sum_{j\neq k} p_j|a_{kj}| = \frac{1}{1+\varepsilon}R_k < |a_{kk}|,$$

$$\frac{1}{p_i}\sum_{j\neq i} p_j|a_{ij}| = R_i + \varepsilon|a_{ik}|,\ i \neq k,\ \forall\varepsilon > 0.$$

Так как $R_i < |a_{ii}|$, то можно выбрать такое малое $\varepsilon > 0$, что для всех $i \neq k$ будет верно $R_i + \varepsilon|a_{ik}| < |a_{ii}|$. Таким образом, $|a_{ii}| > \frac{1}{p_i}\sum\limits_{j\neq i} p_j|a_{ij}| \quad \forall i = \overline{1,n}$, т.е. точка $z = 0$ внешняя для области $G(D^{-1}AD)$ при $D = \mathrm{diag}\{p_1, \ldots, p_n\}$, а значит, A обратима. □

Пример 8.14 (Диагональное преобладание).

Рассмотрим матрицу

$$A = \begin{bmatrix} 7 & -3 & 2 & 0 \\ 2 & 5 & 1 & -1 \\ 0 & 0 & -3 & -2 \\ 1 & 0 & -1 & 2 \end{bmatrix}.$$

Так как все диагональные элементы матрицы A ненулевые и она является матрицей с диагональным преобладанием:

$$\begin{aligned} |7| &> |-3| + |2| = 5, \\ |5| &> |2| + |1| + |-1| = 4, \\ |-3| &> |-2| = 2, \\ |2| &= |1| + |-1| = 2, \end{aligned}$$

причем для $i = 1, 2, 3$ имеют место условия Адамара, то матрица A обратима, т.е. $\det A \neq 0$. (Проверка: $\det A = -363 \neq 0$).

Далее будет сформулирован результат об обратимости с еще более слабыми требованиями строгого диагонального преобладания. Для этого, однако, нам понадобится свойство неразложимости матрицы. Рассмотрим это свойство.

Определение 8.2. Матрица $A \in \mathbb{C}^{n\times n}$ называется *разложимой,* если перестановкой строк вместе с одноименной перестановкой столбцов она может быть приведена к виду

$$\tilde{A} = \begin{bmatrix} B & C \\ 0 & D \end{bmatrix}, \tag{8.23}$$

где B и D — квадратные матрицы. В противном случае матрица A называется *неразложимой.*

Упражнение 8.4. Доказать, что разложимость матрицы A означает, что существует такая матрица перестановок $P \in \mathbb{C}^{n\times n}$ и некоторое целое число $1 \le r \le n-1$, такие, что

$$P^T AP = \begin{bmatrix} B & C \\ 0 & D \end{bmatrix}, \tag{8.24}$$

где $B \in \mathbb{C}^{r\times r}$, $D \in \mathbb{C}^{(n-r)\times(n-r)}$, $0 \in \mathbb{C}^{(n-r)\times r}$ — нулевая матрица.

Существуют различные способы проверки матрицы на неразложимость. Укажем некоторые из них.

Обозначим

$$|A| \equiv [|a_{ij}|]$$

— $(n \times n)$-матрицу абсолютных значений элементов матрицы $A \in \mathbb{C}^{n\times n}$

$$M(A) \equiv [\mu_{ij}]$$

— $n \times n$,-матрицу с элементами μ_{ij} где

$$\begin{aligned} &\mu_{ij} = 1, \text{ если } a_{ij} \ne 0, \\ &\mu_{ij} = 0, \text{ если } a_{ij} = 0. \end{aligned} \tag{8.25}$$

Определение 8.3. Матрица $M(A)$ с элементами (8.25) называется *индикаторной матрицей* для A.

Поскольку разложимость матрицы связана только с расположением ненулевых внедиагональных элементов матрицы A, диагональные элементы и точные значения внедиагональных элементов несущественны. Поэтому верно

Замечание 8.1. Все три матрицы A, $|A|$, $M(A)$ являются разложимыми или неразложимыми одновременно.

Для матрицы $A = [a_{ij}]$ запись $A \ge 0$ $(A > 0)$ означает, что все элементы a_{ij} неотрицательны (положительны).

Теорема 8.11. *Матрица $A \in \mathbb{C}^{n\times n}$ тогда и только тогда неразложима, когда $(E_n + |A|)^{n-1} > 0$ или, что эквивалентно, $[E_n + M(A)]^{n-1} > 0$.*

Доказательство. Рассмотрим вектор $x \ge 0$, $x \ne 0$, и положим

$$y = (E_n + |A|)x = x + |A|x. \tag{8.26}$$

Так как $|A| \ge 0$, то $|A|x \ge 0$ и y имеет по крайней мере столько же ненулевых (и, следовательно, положительных) элементов, сколько и x. Если y еще не положительный, то докажем, что y будет иметь по крайней мере на один ненулевой элемент больше, чем x.

Предположим противное. Пусть ненулевые компоненты вектора x стоят первыми. Тогда можно записать:

$$x = \begin{bmatrix} u \\ 0 \end{bmatrix}, \quad y = \begin{bmatrix} v \\ 0 \end{bmatrix}, \tag{8.27}$$

где $u > 0$, $\quad v > 0$. Разобьем $|A|$ соответственно на блоки

$$|A| = \begin{bmatrix} |A_{11}| & |A_{12}| \\ |A_{21}| & |A_{22}| \end{bmatrix},$$

где $|A_{11}|$, $|A_{22}|$— квадратные матрицы, и рассмотрим (8.26) с учетом (8.27):

$$y = \begin{bmatrix} u \\ 0 \end{bmatrix} + \begin{bmatrix} |A_{11}| & |A_{12}| \\ |A_{21}| & |A_{22}| \end{bmatrix} \begin{bmatrix} u \\ 0 \end{bmatrix} = \begin{bmatrix} u + |A_{11}|u \\ |A_{21}|u \end{bmatrix} = \begin{bmatrix} v \\ 0 \end{bmatrix},$$

следовательно, $|A_{21}|u = 0$.

Так как $u > 0$ и $|A_{21}|u = 0$, то отсюда следует $A_{21} = 0$, что противоречит предположению о неразложимости матрицы A.

Повторяя это рассуждение $n-1$ раз, получаем, что для произвольного ненулевого $x \ge 0$ будет $(E_n + |A|)^{n-1}x > 0$. Полагая $x = e_j$, $\quad j = \overline{1,n}$, убеждаемся в том, что $(E_n + |A|)^{n-1} > 0$. Окончание доказательства следует из замечания 8.1. □

Упражнение 8.5. Если A разложима, то доказать, что при любом положительном целом m матрица A^m также будет разложимой.

Пример 8.15 (Неразложимая матрица). Для матрицы

$$A = \begin{bmatrix} 0 & -3 & 0 \\ 1 & 0 & -1 \\ 0 & 2 & 0 \end{bmatrix}$$

имеем

$$|A| = \begin{bmatrix} 0 & 3 & 0 \\ 1 & 0 & 1 \\ 0 & 2 & 0 \end{bmatrix}, \quad M(A) = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix},$$

$$[E_3 + M(A)]^2 = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}^2 = \begin{bmatrix} 2 & 2 & 1 \\ 2 & 3 & 2 \\ 1 & 2 & 2 \end{bmatrix} > 0,$$

значит, матрица A неразложима.

Пример 8.16 (Разложимая матрица). Для матрицы

$$A = \begin{bmatrix} 4 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

имеем,

$$M(A) = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}.$$

Непосредственным вычислением находим

$$[E_3 + M(A)]^2 = \begin{bmatrix} 2 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 1 & 2 \end{bmatrix}^2 = \begin{bmatrix} 1 & 3 & 3 \\ 0 & 2 & 2 \\ 0 & 2 & 2 \end{bmatrix} \geq 0.$$

Значит, матрица A разложима. Действительно, она уже представлена в виде (8.23), где $B = [4]$, $C = [2, 1]$, $D = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$, т.е. разложена.

С точки зрения неразложимости важны не значения элементов матрицы, а только расположение ее ненулевых элементов.

Расположение ненулевых элементов матрицы A можно сделать наглядным, пользуясь ассоциированным с A графом.

Ориентированный граф матрицы $A \in \mathbb{C}^{n\times n}$ обозначается $\Gamma(A)$ и представляет собой ориентированный граф с n узлами $P_1, P_2, \ldots, P_n$, в котором дуга из P_i в P_j присутствует тогда и только тогда, когда $a_{ij} \neq 0$ (или $\mu_{ij} \neq 0$).

Пример 8.17 (Орграф матрицы). Для матриц

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$$

ориентированные графы изображены на рисунке 8.14 (a), (b) и (c), соответственно.

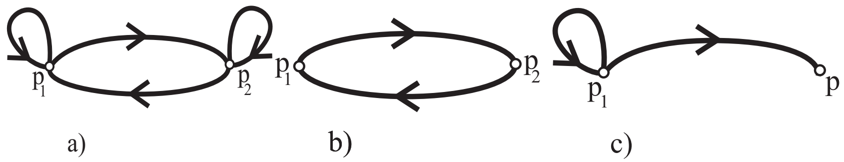


Рис. 8.14

Для матриц из примеров 8.15 и 8.16 их ориентированные графы имеют вид, представленный на рисунке 8.15 (a) и (b) соответственно.

Определение 8.4. *Ориентированным путем* в графе Γ из вершины P_{i_1} в вершину P_{i_k} называется последовательность дуг $(P_{i_1}P_{i_2}), (P_{i_2}P_{i_3}), \ldots, (P_{i_{k-1}}P_{i_k})$. *Длина ориентированного пути* — это число дуг в нем, если оно конечно.

Существование пути из P_{i_1} в P_{i_k} в орграфе $\Gamma(A)$ эквивалентно существованию последовательности ненулевых элементов $a_{i_1i_2}, a_{i_2i_3}, \ldots, a_{i_{k-1}i_k}$ в матрице A.

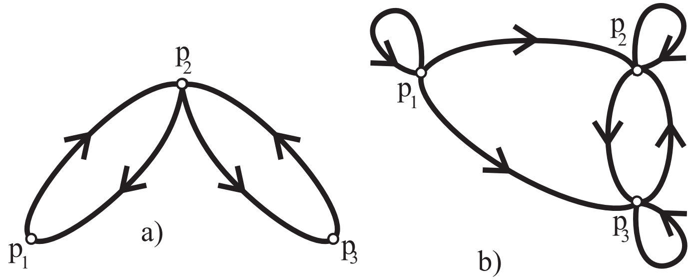


Рис. 8.15

Пример 8.18 (Путь в орграфе матрицы).

Для матрицы A из примера 8.15 и ее орграфа на рис. 8.15 a) путь $(P_1, P_2), (P_2, P_3)$ из вершины P_1 в вершину P_3 сответствует последовательности ненулевых элементов $a_{12} = -3$, $a_{23} = -1$.

Определение 8.5. Ориентированный граф называется *сильно связным*, если в нем любые два узла P_i, P_j соединены ориентированным путем конечной длины, начинающимся в P_i и заканчивающимся в P_j.

Теорема 8.12. *[10, стр. 433] Матрица $A \in \mathbb{C}^{n\times n}$ тогда и только тогда неразложима, когда ориентированный граф $\Gamma(A)$ сильносвязный.*

Следствие 8.2. *Матрица, у которой все элементы ненулевые, неразложима.*

Пример 8.19 (Связность графа и неразложимость).

В примере 8.17 сильно связными являются графы $\Gamma(A)$ и $\Gamma(B)$. Граф $\Gamma(C)$ не является сильно связным, так как в нем нет пути из P_2 в P_1. Следовательно, матрицы A и B — неразложимы, а матрица C — разложима.

Граф матрицы из примера 8.15 является сильно связным, а граф матрицы из примера 8.16 не является сильно связным,

так как в нем нет путей из P_2 в P_1 и из P_3 в P_1 (см. рис. 8.15). Следовательно, матрица из примера 8.15 неразложима, а матрица из примера 8.16 — разложима (сравни с выводами в примерах 8.15 и 8.16).

Теорема 8.13 (Ольги Тауски). *[9, стр. 384] Если для неразложимой матрицы A выполняются ослабленные условия Адамара (8.22) и по крайней мере в одном из этих условий имеет место знак строгого неравенства ($>$), то матрица A — невырожденная.*

Пример 8.20 (Невырожденная матрица). Для матрицы

$$A = \begin{bmatrix} 2 & -1 & 1 \\ -1 & 5 & -2 \\ 1 & 3 & -4 \end{bmatrix}$$

выполнены условия теоремы Ольги Тауски. Действительно, так как все элементы матрицы A ненулевые, то она неразложима. Кроме того, выполнены ослабленные условия Адамара $(|2| = |-1| + |1| = 2,\quad |5| > |-1| + |-2| = 3,\ |-4| = |1| + |3| = 4)$, причем одно из них — в сильной форме. Поэтому матрица A обратима. Непосредственное вычисление показывает, что $\det A = -30 \neq 0$, т.е. матрица A обратима.

Задачи и упражнения

1. Не вычисляя собственные значения матрицы A, получить максимум информации о расположении собственных значений и величине спектрального радиуса матрицы A:

a) $\begin{bmatrix} 0 & 2 \\ 3 & 2i \end{bmatrix}$ b) $\begin{bmatrix} 1 & 2 \\ -1 & 4 \end{bmatrix}$ c) $\begin{bmatrix} i & -2 \\ 3 & 1 \end{bmatrix}$ d) $\begin{bmatrix} -1 & 2 \\ i & 2i \end{bmatrix}$

e) $\begin{bmatrix} 7 & -16 & 8 \\ -16 & 7 & -8 \\ 8 & -8 & -5 \end{bmatrix}$ f) $\begin{bmatrix} 0 & 1 & 0 & i \\ 1 & 6 & 1 & 1 \\ \frac{i}{2} & i & 5 & 0 \\ 0 & \frac{i}{2} & \frac{1}{2} & -2 \end{bmatrix}$.

2. Для матриц $A \in \mathbb{C}^{n\times n}$, приведенных в задаче 1, рассмотреть матрицы вида $D^{-1}AD$, где $D = \mathrm{diag}(p_1, \ldots, p_n)$, и выяснить, можно ли улучшить локализацию собственных значений.

3. Вычислить собственные значения матриц A, приведенных в задаче 1, и прокомментировать качество полученных там оценок.

4. Каким образом оценка спектрального радиуса (8.3) связана с оценкой (7.15)?

5. Показать, что каждое собственное значение матрицы A принадлежит множеству $\bigcap\limits_D G\left(D^{-1}AD\right)$, где пересечение берется по всем диагональным матрицам с положительными диагональными элементами.

6. Пусть a, b, c, d — положительные числа и $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

   (a). Найти диагональную матрицу $\tilde{D}$ такую, что
   $$\left\|\tilde{D}^{-1}A\tilde{D}\right\|_\infty = \min_D \left\|D^{-1}AD\right\|_\infty.$$
   Минимум берется по всем диагональным матрицам D с положительными диагональными элементами.

   (b). Вычислить $\left\|\tilde{D}^{-1}A\tilde{D}\right\|_\infty \equiv r$.

   (c). Вычислить $\rho(A)$.

   (d). Сравнить r и $\rho(A)$.

7. Для матрицы
   $$A = \begin{bmatrix} 1 & 1 \\ -1,5 & 2 \end{bmatrix}$$
   (a) найти $\rho(A)$; (b) показать, что $\rho(A) < \min\limits_D \left\|D^{-1}AD\right\|_\infty$, где минимум берется по всем $D = \mathrm{diag}(p_1, p_2)$, $\quad p_1 \geq 0, p_2 \geq 0$.

8. Доказать следующие оценки спектрального радиуса:

   (a) $\rho(A) \leq \min\limits_{p_1,\ldots,p_n>0} \max\limits_{1\leq i\leq n} \frac{1}{p_i} \sum\limits_{j=1}^{n} p_j|a_{ij}|$;

   (b) $\rho(A) \leq \min\limits_{p_1,\ldots,p_n>0} \max\limits_{1\leq i\leq n} p_j \sum\limits_{i=1}^{n} \frac{1}{p_i}|a_{ij}|$.

9. С помощью матриц $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$, $\begin{bmatrix} 1 & 1 \\ 1-\varepsilon & 1 \end{bmatrix}$ показать, что простое диагональное преобладание не гарантирует обратимость, а строгое диагональное преобладание не является ее необходимым условием.

10. Для матрицы $A = \begin{bmatrix} 1 & 4 \\ 1 & 6 \end{bmatrix}$ сравнить строчную и столбцовую области локализации Гершгорина с областью Островского для $\alpha = \frac{1}{2}$. Какую оценку дает теорема Островского для спектрального радиуса матрицы A? Сравнить ее с оценками Гершгорина.

11. Вывести условия невырожденности матрицы из теорем a) Островского и b) Брауэра.

12. Используя теорему Островского и параметры $p_1, \ldots, p_n$, получить области локализации собственных значений вида (8.4),(8.5).

13. Объяснить, почему ноль не является собственным значением матрицы
    $$A = \begin{bmatrix} n & 1 & \ldots & 1 \\ 1 & n & \ldots & 1 \\ \ldots & \ldots & \ldots & \ldots \\ 1 & 1 & \ldots & n \end{bmatrix} \in \mathbb{R}^{n\times n}$$
    и, следовательно, матрица A невырожденная.

14. Используя теорему Гирша, дать оценки собственных значений матриц из задачи 1 и сравнить с результатами задачи 3.

15. Для матрицы
    $$A = \begin{bmatrix} 0 & 1 & 0 & i \\ 1 & 6 & 1 & 1 \\ \frac{i}{2} & i & 5 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & -2 \end{bmatrix}$$
    получить, что A имеет собственное значение в круге $|z+2| \leq \frac{1}{2}$.

16. Пусть $A \in \mathbb{R}^{n\times n}$ — вещественная матрица, все круги Гершгорина которой попарно не пересекаются. Показать, что все ее собственные значения вещественны.

17. Пусть $A \in \mathbb{R}^{n\times n}$ — комплексная матрица с вещественными диагональными элементами и характеристическим многочленом, имеющим только вещественные коэффициенты. Показать, что если все круги Гершгорина не пересекаются, то все собственные значения вещественны.

18. С помощью каких условий можно установить невырожденность матрицы
$$A = \begin{bmatrix} -2 & 1 & 0 \\ 1 & 3 & 1 \\ 2 & 2 & 4 \end{bmatrix}?$$

19. Используя теорему Гершгорина и принадлежность матрицы специальному типу (эрмитова, косоэрмитова, унитарная), указать область локализации собственных значений следующих матриц:
a) $\begin{bmatrix} \frac{1}{3} & -\frac{2}{3} & -\frac{2}{3} \\ -\frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \\ -\frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \end{bmatrix}$, b) $\begin{bmatrix} 1 & 0 & 1-i \\ 0 & 2 & -i \\ 1+i & i & 3 \end{bmatrix}$,
c) $\begin{bmatrix} i & 0 & 1-i \\ 0 & 2i & -1 \\ -1-i & 1 & 3i \end{bmatrix}$, d) $\begin{bmatrix} 2i & 1-i & i \\ -1-i & 0 & 4 \\ i & -4 & 2i \end{bmatrix}$,
e) $\begin{bmatrix} -2 & 1+i & -1 \\ 1-i & 0 & 4i \\ -1 & -4i & -2 \end{bmatrix}$.

20. Пусть дана матрица $A \in \mathbb{R}^{n\times n}$ и в результате перестановки одноименных строк и столбцов i, j из нее получена матрица $\tilde{A}$. Как изменится при этом ориентированный граф $\Gamma(A)$? Повлияет ли это на его связность?

21. Пусть $A \in \mathbb{R}^{n\times n}$ — матрица со строгим диагональным преобладанием. Показать, что
$$|a_{kk}| > C_k = \sum_{i\neq k} |a_{ik}|$$
хотя бы для одного значения $k = 1, \ldots, n$.

22. Пусть матрица $A \in \mathbb{R}^{n\times n}$ *идемпотентна* (т.е. $A^2 = A$), но $A \neq E_n$. Показать, что A не может быть матрицей со строгим диагональным преобладанием.

23. Доказать, что неразложимая матрица не может иметь нулевых строк и столбцов.

24. Пусть матрица $A \in \mathbb{R}^{n\times n}$ имеет строгое диагональное преобладание и пусть $D = \mathrm{diag}(a_{11}, a_{22}, \ldots, a_{nn})$. Показать, что D обратима и $\rho(E_n - D^{-1}A) < 1$.

25. Получить условия сильной связности орграфа матрицы $A = [a_{ij}]_{k\times k}$, $k = 2, 3$, выраженные через ее элементы.

26. Доказать, что приведенное ниже определение разложимой матрицы эквивалентно введенному на стр. 281 определению.

Определение 8.6. Квадратная матрица A порядка n называется *разложимой,* если существует такое разбиение индексов $J = \{1, 2, \ldots, n\}$ на два непересекающиеся подмножества $S = \{i_1, \ldots, i_k\}$, $S' = \{j_1, \ldots, j_m\}$, $J = S \bigcup S'$, $S = \bigcap S' = \emptyset$ $(k + m = n)$, что $a_{ij} = 0$, $i \in S$, $j \in S'$. В противном случае матрицу A называют *неразложимой.*

27. Доказать, что матрица $A \in \mathbb{C}^{n\times n}$ будет невырожденной, если в каждой ее строке имеется доминирующий (не обязательно диагональный!) элемент и эти n доминирующих элементов расположены в различных столбцах.

28. Показать, что граф Γ сильносвязен, если любые два его узла принадлежат хотя бы одному общему циклу, и что обратное утверждение неверно.

29. Доказать, что если A — матрица со строгим диагональным преобладанием и ее диагональные элементы положительны, то все ее собственные значения имеют положительную вещественную часть.

30. Доказать, что если $A \in \mathbb{C}^{n\times n}$ — эрмитова матрица со строгим диагональным преобладанием и положительными диагональными элементами, то A положительно определена.

31. Доказать, что если A — эрмитова матрица со строгим диагональным преобладанием и все ее диагональные элементы положительны, то все собственные значения матрицы положительны.

Вопросы для повторения

1. Что такое строчная область Гершгорина?
2. Что такое столбцовая область Гершгорина?
3. Для каких матриц справедлива теорема Бендиксона? Сформулируйте ее.
4. Какую оценку собственных значений матрицы дает неравенство Шура?
5. Когда в неравенстве Шура имеет место равенство?
6. Какую оценку спектрального радиуса можно получить из теоремы Гершгорина? Островского? Брауэра?
7. Какую оценку собственных значений матрицы дает теорема Брауна?
8. Что такое матрица с диагональным преобладанием?
9. Какая матрица называется матрицей со строгим диагональным преобладанием?
10. Где на комплексной плоскости расположены собственные значения матрицы со строгим диагональным преобладанием, все диагональные элементы которой положительны?
11. Какие условия обратимости матрицы следуют из теоремы Гершгорина? Островского? Брауэра?

12*. Сформулируйте столбцовую версию теоремы Брауэра.

13. Что можно сказать о расположении собственных значений эрмитовой матрицы? Косоэрмитовой матрицы? Унитарной матрицы? Ортогональной матрицы?
14. Как выглядит столбцовый вариант условий Адамара?
15. Какая матрица называется разложимой? Неразложимой?
16. Сформулируйте столбцовую версию теоремы Ольги Тауски.
17. Что такое индикаторная матрица для матрицы A?

18. Сформулируйте критерии неразложимости матрицы.
19. Что такое ориентированный граф матрицы A?
20. Какой граф называется сильно связным?

8.2*. Возмущения собственных значений

Изучим вопрос о том, как изменяются собственные значения при небольшом изменении («возмущении») элементов матрицы. Эта проблема связана с неизбежными погрешностями, с которыми нам известна исходная информация в любой практической задаче и со столь же неизбежными погрешностями округления в процессе машинного вычисления.

Пусть $\lambda_i,\ i = \overline{1,n}$ — собственные значения матрицы A, а $\tilde{\lambda}_i,\ i = \overline{1,n}$ — собственные значения возмущенной матрицы $A + \delta A$. Поскольку собственные значения непрерывно зависят от элементов матрицы, то при «достаточно малой» матрице возмущений δA собственные значения $\tilde{\lambda}_i$ не сильно отличаются от λ_i. Укажем более точные оценки для некоторых типов матриц.

Теорема 8.14. *Пусть* $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ — *диагональная матрица,* $\delta\Lambda = [l_{ij}] \in \mathbb{C}^{n\times n}$ — *произвольная матрица. Если* $\tilde{\lambda}$ — *собственное значение матрицы* $\Lambda + \delta\Lambda$*, то найдется такое собственное значение* λ_i *матрицы* Λ*, что*

$$\left|\tilde{\lambda} - \lambda_i\right| \le \|\delta\Lambda\|_\infty . \tag{8.28}$$

Доказательство. Из теоремы Гершгорина для возмущенной матрицы $\Lambda + \delta\Lambda$ следует, что

$$\sigma(\Lambda + \delta\Lambda) \subseteq \bigcup_{i=1}^{n} \Big\{ z \in \mathbb{C} : |z - \lambda_i - l_{ii}| \le \sum_{\substack{j=1\\ j\ne i}}^{n} |l_{ij}| \Big\} = G(\Lambda + \delta\Lambda).$$

Из свойств функции модуль получим:

$$|z-\lambda_i| - |l_{ii}| \le |z-\lambda_i - l_{ii}| \le \sum_{j\ne i}^{n} |l_{ij}|.$$

Значит, для точек $z \in G(\Lambda + \delta\Lambda)$ справедливо

$$|z-\lambda_i| \le \sum_{j=1}^{n} |l_{ij}| \le \max_{1\le i\le n} \sum_{j=1}^{n} |l_{ij}| = \|\delta\Lambda\|_\infty .$$

Следовательно, построенная выше область Гершгорина, а значит, и спектр матрицы $\Lambda + \delta\Lambda$, лежат в области

$$\bigcup_{i=1}^{n} \{z \in \mathbb{C} : |z-\lambda_i| \le \|\delta\Lambda\|_\infty\} . \qquad \square$$

Напомним, что матрица $A \in \mathbb{C}^{n\times n}$ называется *диагонализуемой*, если она подобна диагональной матрице, т.е. существует невырожденная $n \times n$-матрица S и диагональная матрица $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ такие, что $A = S\Lambda S^{-1}$.

Теорема 8.15. *Пусть $A \in \mathbb{C}^{n\times n}$ — диагонализуемая матрица $A = S\Lambda S^{-1}$, $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, $\delta A \in \mathbb{C}^{n\times n}$. Если $\tilde\lambda$ — собственное значение матрицы $A+\delta A$, то найдется собственное значение λ_i матрицы A, для которого*

$$|\tilde\lambda - \lambda_i| \le \mathrm{cond}_\infty(S)\, \|\delta A\|_\infty . \tag{8.29}$$

Если A — нормальная матрица, то для некоторого ее собственного значения λ_i

$$\left|\tilde\lambda - \lambda_i\right| \le \|\delta A\|_2 . \tag{8.30}$$

Таким образом, нормальные матрицы идеально обусловлены по отношению к вычислению собственных значений.

Замечание. В [10, стр. 436] доказано, что оценка (8.29) верна также для всех матричных норм $\|\cdot\|$ таких, что

$$\|D\| = \max_{1\le i\le n} |d_i|$$

для всех диагональных матриц $D = \mathrm{diag}(d_1, \ldots, d_n)$.

Упражнение 8.6. Доказать теорему 8.15, используя $\sigma(A+\delta A) = \sigma(\Lambda + S^{-1}\delta A S)$, кольцевое свойство матричной нормы, определение числа обусловленности и спектральную теорему для нормальных матриц.

Заметим, что столбцы матрицы S из (8.29) являются собственными векторами матрицы A. Таким образом, возмущения в собственных значениях диагонализуемой матрицы A прямопропорциональны числу обусловленности матрицы ее собственных векторов.

Пример 8.21 (Возмущение спектра нормальной матрицы).

Пусть $A \in \mathbb{C}^{3\times 3}$ — заданная эрмитова (а значит, и нормальная) матрица, и пусть каким-то способом найдена унитарная матрица U, такая, что

$$\bar A = UAU^* = \begin{bmatrix} 3,05 & -0,06 & 0,02 \\ -0,06 & -6,91 & 0,07 \\ 0,02 & 0,07 & 8,44 \end{bmatrix}.$$

Таким образом, матрица A унитарно подобна матрице $\bar A$, а значит, имеет те же самые собственные значения. Для оценки собственных значений матрицы $\bar A$ представим ее в виде суммы $A = \tilde A + \delta A$, где

$$\tilde A = \mathrm{diag}(3; -7; 8,5), \qquad \delta A = \begin{bmatrix} 0,05 & -0,06 & 0,02 \\ -0,06 & 0,09 & 0,07 \\ 0,02 & 0,07 & -0,06 \end{bmatrix}.$$

Вычислив сингулярные числа матрицы δA :

$$\sigma_1 \approx 0,14; \quad \sigma_2 \approx 0,1; \quad \sigma_3 \approx 0,035,$$

находим $\|\delta A\|_2 \approx 0,14$. Из (8.30) имеем, что для собственных значений λ_i, $i = \overline{1,3}$, матрицы $\bar{A}$, а, значит и A, справедливы оценки

$$|\lambda_1 - 3| \leq 0,14; \ |\lambda_2 + 7| \leq 0,14; \ |\lambda_3 - 8,5| \leq 0,14.$$

Непосредственное вычисление дает собственные значения матрицы $\bar{A}$:

$$\lambda_1 = 3,05029; \ \lambda_2 = -6,91068; \ \lambda_3 = 8,44039,$$

что подтверждает полученные оценки.

Пример 8.22 (Возмущение спектра недиагонализуемой матрицы). Рассмотрим, матрицы

$$A = \begin{bmatrix} 0 & a & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \ \delta A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & \varepsilon \\ \varepsilon & 0 & 0 \end{bmatrix}, \ a, \varepsilon \geq 0.$$

Т.к. $AA^* \neq A^*A$, то A не является нормальной матрицей. Все собственные значения матрицы A равны нулю. Собственными значениями матрицы $A + \delta A$ будут три различных значения $\sqrt[3]{a\varepsilon^2}$. Каково бы ни было $\varepsilon > 0$, все собственные значения матрицы $A + \delta A$ можно сделать сколь угодно большими, подбирая соответствующее a. Если бы матрица A была нормальной, то согласно (8.4) собственные значения матрицы $A + \delta A$ были бы не больше $\|\delta A\|_2 = \varepsilon$.

В отличие от собственных значений собственные векторы даже диагонализуемой матрицы могут претерпевать радикальные изменения при очень малых возмущениях элементов матриц.

Пример 8.23 (Возмущение собственных векторов).

Рассмотрим матрицы

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \delta A = \begin{bmatrix} \varepsilon & \delta \\ 0 & 0 \end{bmatrix}.$$

Собственными значениями матрицы $A + \delta A$ являются $\tilde{\lambda}_1 = 1$ и $\tilde{\lambda}_2 = 1 + \varepsilon$; при $\varepsilon\delta \neq 0$ соответствующими нормированными собственными векторами будут

$$\frac{1}{(\varepsilon^2 + \delta^2)^{\frac{1}{2}}} \begin{bmatrix} -\delta \\ \varepsilon \end{bmatrix}, \ \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

При надлежащем выборе отношения $\frac{\varepsilon}{\delta}$ первому собственному вектору можно придать любое желаемое направление, как бы малы ни были ε и δ по отдельности.

Если положить $\varepsilon = 0$, то у возмущенной матрицы $A + \delta A$ при любом $\delta \neq 0$ будет только один, с точностью до скалярного множителя, собственный вектор; сама же A имеет два линейно независимых собственных вектора.

Пусть для матрицы A известен «приближенный собственный вектор» $\hat{x} \neq 0$ и «приближенное собственное значение» $\hat{\lambda}$. Для оценки расстояния от $\hat{\lambda}$ до точного собственного значения матрицы A можно использовать *вектор невязки* $r = A\hat{x} - \hat{\lambda}\hat{x}$.

Теорема 8.16. *Пусть $\|\cdot\|_V$ — векторная норма на $\mathbb{C}^n$; $\|\cdot\|_M$ — матричная норма на $\mathbb{C}^{n\times n}$, согласованная с векторной нормой $\|\cdot\|_V$; $\|D\|_M = \max\limits_{1\leq i\leq n} |d_i|$ для любой диагональной матрицы $D = \operatorname{diag}(d_1, \ldots, d_n)$. Пусть $A \in \mathbb{C}^{n\times n}$ — диагонализуемая матрица, $A = S\Lambda S^{-1}$, $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$. Тогда найдется собственное значение λ_i матрицы A, для которого*

$$|\hat{\lambda} - \lambda_i| \leq \operatorname{cond}_{\|\cdot\|_M}(S) \frac{\|r\|}{\|\hat{x}\|}. \qquad (8.31)$$

Доказательство. Из $A = S\Lambda S^{-1}$ следует

$$r = A\hat{x} - \hat{\lambda}x = S(\Lambda - \hat{\lambda}E)S^{-1}\hat{x}.$$

Пусть $\hat{\lambda}$ не совпадает ни с одним из собственных значений матрицы A. Тогда матрица $(\Lambda - \hat{\lambda}E)$ — невырожденная, и

$$\hat{x} = S(\Lambda - \hat{\lambda}E)^{-1}S^{-1}r.$$

Далее для согласованных норм с учетом кольцевого свойства матричной нормы имеем

$$\begin{aligned}
&\|\hat{x}\| = \left\|S(\Lambda - \hat{\lambda}E)^{-1}S^{-1}r\right\| \le \left\|S(\Lambda - \hat{\lambda}E)^{-1}S^{-1}\right\| \cdot \|r\| \le \\
&\le \|S\| \cdot \left\|S^{-1}\right\| \cdot \left\|(\Lambda - \hat{\lambda}E)^{-1}\right\| \cdot \|r\| = \\
&= \operatorname{cond}_{\|.\|}(S) \left\|(\Lambda - \hat{\lambda}E)^{-1}\right\| \cdot \|r\| = \\
&= \operatorname{cond}_{\|.\|}(S) \min_{1\le i\le n} |\lambda_i - \hat{\lambda}|^{-1} \|r\|.
\end{aligned}$$

Таким образом,

$$\min_{1\le i\le n} |\lambda_i - \hat{\lambda}| \le \operatorname{cond}_{\|.\|}(S)\frac{\|r\|}{\|\hat{x}\|}.$$

Ясно, что неравенство (8.31) будет выполнено и в случае, когда $\hat{\lambda} = \lambda_i$. □

Пример 8.24 (Возмущение спектра). Пусть

$$A = \begin{bmatrix} 1 & \varepsilon \\ \varepsilon & 1 \end{bmatrix},\ \varepsilon > 0.$$

Точные собственные значения матрицы A равны $\lambda_1 = 1 + \varepsilon$ и $\lambda_2 = 1 - \varepsilon$. Если взять $\hat{\lambda} = 1$, $\hat{x} = [1, 0]^T$, то невязка имеет вид $r = [0, \varepsilon]^T$. Так как $\|r\|_2 = \varepsilon$, $\|\hat{x}\|_2 = 1$, то согласно оценке

(8.22) $|\hat{\lambda} - \lambda_i| \le \varepsilon$. Непосредственным вычислением получаем $|\hat{\lambda} - \lambda_1| = |\hat{\lambda} - \lambda_2| = \varepsilon$, что подтверждает полученную оценку.

Малость невязки, однако, не гарантирует, что приближенный собственный вектор будет близок к точному. Собственными векторами матрицы A для любого $\varepsilon > 0$ будут $[1, 1]^T$ и $[1, -1]^T$. Вектор $\hat{x}$ не будет приблизительно параллелен ни одному из них, как бы мало ни было ε.

Задачи и упражнения

1. Доказать, что для симметричной матрицы проблема нахождения собственных значений всегда хорошо обусловлена.

2. Доказать, что нормальные матрицы идеально обусловлены по отношению к вычислению собственных значений.

3. Доказать, что если $D = \operatorname{diag}(d_1, \ldots, d_n)$ — диагональная матрица, а $\delta D \in \mathbb{C}^{n\times n}$ — произвольная, то собственные значения матрицы $D + \delta D$ принадлежат объединению кругов $|z - d_i| \le \|\delta D\|_\infty$.

4. Доказать, что если A — нормальная матрица, то для некоторого ее собственного значения λ_i

$$|\tilde{\lambda} - \lambda_i| \le \frac{\|r\|_2}{\|\tilde{x}\|_2}. \qquad (8.32)$$

Для нормальной матрицы A неравенство (8.32) означает, что если невязка, соответствующая приближениям к собственному значению и собственному вектору, мала, то абсолютная погрешность приближенного собственного значения гарантированно будет малой.

5. Для заданной нормальной матрицы $\bar{A}$ из примера 8.21 известен приближенный собственный вектор $\hat{x} = (-0,004; -0,005; -0,99)^T$ и соответствующее приближенное собственное значение $\hat{\lambda} = 8,5$. Найти вектор невязки r, точный собственный вектор и оценить его отклонение от приближенного и подтвердить оценку (8.31).

6. (*Теорема о возмущениях собственных значений эрмитовой матрицы*). Пусть матрицы $A, B \in \mathbb{C}^{n\times n}$ эрмитовы. Доказать, что $|\lambda_k\,(A+B) - \lambda_k\,(A)\,| \le \rho(B)$, где $\rho(B)$ — спектральный радиус матрицы B.

Вопросы для повторения

1. Как можно оценить возмущения собственных значений диагонализуемой матрицы? Нормальной матрицы?
2. Число обусловленности какой матрицы характеризует величину изменений собственных значений диагонализуемой матрицы?
3. Какие матричные нормы используются для оценки возмущений собственных значений диагонализуемой матрицы? Нормальной матрицы?
4. Как можно оценить расстояние от «приближенного» до точного собственного значения диагонализуемой матрицы?
5. Как вычисляется вектор невязки для оценки приближенного вычисленного собственного значения?
6. Какую оценку можно дать приближенному собственному значению диагонализуемой матрицы? Нормальной матрицы?
7. Гарантирует ли малость невязки приближенного собственного значения близость приближенного собственного вектора к точному?

Необходимо усвоить.

Основные понятия и теоремы

- орграф матрицы
- сильно связный орграф
- круги Гершгорина

- матрица
  - – – * индикаторная
  - – – разложимая
  - – – неразложимая
  - – – перестановок
  - – – с диагональным преобладанием
  - – – со строгим диагональным преобладанием
  - – – диагонализуемая
- область Гершгорина
  - – – столбцовая
  - – – строчная
- окружности Гершгорина
- условия Адамара (строгие, нестрогие)
- теоремы Гершгорина
- * теорема Ольги Тауски
- критерии и условия неразложимости матрицы
- неравенство Шура
- теорема Бендиксона
- теорема Гирша
- * теорема Островского
- теорема Брауна
- * теорема Брауэра
- теорема Адамара
- * теоремы об оценках собственных значений возмущенной диагональной матрицы
- * теоремы об оценках собственных значений возмущенной диагонализуемой матрицы

Основные умения и навыки

- строить области локализации собственных значений заданной матрицы (по теоремам Гершгорина, * Островского)
- уточнять области локализации специальных типов матриц (эрмитовых, унитарных) на основе их свойств
- анализировать свойства матрицы (сходимость, невырожденность, устойчивость, эрмитовость, унитарность и т.п.) по областям локализации ее спектра
- оценивать величину спектрального радиуса заданной матрицы с использованием областей локализации ее спектра
- строить ориентированный граф матрицы и устанавливать его связность
- выяснять, является ли заданная матрица разложимой

Часть IV

НЕОТРИЦАТЕЛЬНЫЕ МАТРИЦЫ

Матрицы с неотрицательными элементами возникают в математической экономике (неотрицательные матрицы моделей экономического взаимодействия, вход-выход анализ Леонтьева), биологии (модели динамики популяций), в теории вероятностей при исследовании цепей Маркова (стохастические матрицы при исследовании марковских процессов принятия решений), в теории малых колебаний упругих систем (осцилляционные матрицы), теории игр, при моделировании физических процессов (каскады космических лучей, деление ядер), медицине и т.д.

Цель изучения

Изучить основные классы неотрицательных матриц и их свойства.

Следует повторить

- определение максимальной строчной и максимальной столбцовой матричной нормы
- связь спектрального радиуса и нормы матрицы
- определение и свойства неразложимой матрицы
- понятие сильносвязного графа
- матрица перестановок и ее свойства
- понятие миноров матрицы
- теорема Шура
- выпуклое компактное множество, его свойства, крайние точки

Глава 9

Положительные и неотрицательные матрицы

Рекомендуемая литература:

[1]*, [3], [5], [7], [9], [10], [13], [15], [16], [19]

9.1. Определения и свойства

Определение 9.1. Матрица $A = [a_{ij}] \in \mathbb{R}^{n\times m}$ называется *неотрицательной (положительной)* и это обозначается как $A \geq 0$ $(A > 0)$, если $a_{ij} \geq 0$ $(a_{ij} > 0)$, $i = \overline{1,n}$, $j = \overline{1,m}$.

Аналогично определяются отношения $\leq$ и $<$ и понятия *неположительной (отрицательной)* матрицы.

Если $A - B \geq 0$ $(A - B > 0)$, то пишут $A \geq B$ $(A > B)$.

По определению считаем $|A| \equiv [|a_{ij}|]$.

Пример 9.1 (Неотрицательные матрицы).

$$A = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix} \geq 0,\ B = \begin{bmatrix} 4 & 3 \\ 1 & 1 \end{bmatrix} > 0,\ |A| = A,\ |B| = B,$$
$$C = \begin{bmatrix} -1 & -3 \\ 0 & 0 \end{bmatrix} \leq 0,\ D = \begin{bmatrix} -1 & -2 \\ -4 & -5 \end{bmatrix} < 0,$$
$$|C| = \begin{vmatrix} 1 & 3 \\ 0 & 0 \end{vmatrix},\ |D| = \begin{bmatrix} 1 & 2 \\ 4 & 5 \end{bmatrix}.$$

Матрица $F = \begin{bmatrix} -1 & 3 \\ 0 & 4 \end{bmatrix}$ не является ни одной из определенных выше.

Свойства неотрицательных матриц

Пусть $A, B \in \mathbb{R}^{n\times m}$.

1^0. $|A| \geq 0\ \forall A$; $|A| = 0 \Leftrightarrow A = 0$.

2^0. $|Ax| \leq |A|\,|x|$.

3^0. $|A^m| \leq |A|^m,\ \forall m = 1, 2, \ldots$

4^0. Если $0 \leq A \leq B$, то $0 \leq A^m \leq B^m,\ \ \forall m = 1, 2, \ldots$.

5^0. Если $A > 0$, $x \geq 0$ и $x \neq 0$, то $Ax > 0$.

6^0. Если $|A| \leq |B|$, то $\|A\|_E \leq \|B\|_E$.

7^0. $\|A\|_E = \|\,|A|\,\|_E$.

Упражнение 9.1. Доказать свойства 1^0— 7^0.

Теорема 9.1. *Пусть $A, B \in \mathbb{R}^{n\times n}$. Если $|A| \leq B$, то*

$$\rho(A) \leq \rho\left(|A|\right) \leq \rho(B).$$

Доказательство. Во-первых, из свойств неотрицательных матриц следует справедливость следующей цепочки неравенств для $\forall\ m = 1, 2, \ldots$:

$$|A^m| \overset{3^0}{\leq} |A|^m \overset{1^0,\,4^0}{\leq} B^m.$$

Далее из этих неравенств имеем для $\forall\ m = 1, 2, \ldots$:

$$\|A^m\|_E \overset{7^0}{\leq} \|\,|A|^m\|_E \overset{6^0}{\leq} \|B^m\|_E \quad \Rightarrow$$
$$\Rightarrow \|A^m\|_E^{\frac{1}{m}} \leq \|\,|A|^m\|_E^{\frac{1}{m}} \leq \|B^m\|_E^{\frac{1}{m}}.$$

Переходя в последних равенствах к пределу при $m \to \infty$ согласно равенству $\rho(A) = \lim\limits_{k\to\infty} \left\|A^k\right\|^{\frac{1}{k}}$ (см. следствие 7.2 стр. 238), получаем соотношения теоремы 9.1. $\square$

Следствие 9.1. *Пусть $A, B \in \mathbb{R}^{n\times n}$. Если $0 \leq A \leq B$, то $\rho(A) \leq \rho(B)$.*

Упражнение 9.2. Доказать следствие 9.1.

Лемма 9.1. *Пусть $A \in \mathbb{R}^{n\times n}$ и $A \geq 0$. Если строчные суммы для A постоянны, то $\rho(A) = \|A\|_\infty$. Если для A постоянны столбцовые суммы, то $\rho(A) = \|A\|_1$.*

Доказательство. Пусть строчные суммы постоянны и равны α. Тогда $\|A\|_\infty = \alpha$ и покажем, что вектор $x = [1, \ldots, 1]^T$ является собственным вектором матрицы A, соответствующим собственному значению $\|A\|_\infty = \alpha$. Действительно,

$$Ax = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} = \begin{bmatrix} \sum\limits_{i=1}^{n} a_{1i} \\ \vdots \\ \sum\limits_{i=1}^{n} a_{ni} \end{bmatrix} =$$

$$= \|A\|_\infty \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} = \|A\|_\infty \, x.$$

Так как $\rho(A) \le \|A\|_\infty$ (см. (7.15)) и существует собственное значение матрицы A, равное $\|A\|_\infty$, то $\rho(A) = \|A\|_\infty$.

Аналогично доказывается $\rho(A) = \|A\|_1$, если рассмотреть матрицу A^T. □

Теорема 9.2. *Пусть* $A \in \mathbb{R}^{n\times n}$ *и* $A \ge 0$. *Тогда*

$$\min_{1\le i\le n} \sum_{j=1}^{n} a_{ij} \le \rho(A) \le \max_{1\le i\le n} \sum_{j=1}^{n} a_{ij} = \|A\|_\infty\,,$$

$$\min_{1\le j\le n} \sum_{i=1}^{n} a_{ij} \le \rho(A) \le \max_{1\le j\le n} \sum_{i=1}^{n} a_{ij} = \|A\|_1\,.$$

Доказательство. Положим

$$\alpha = \min_{1\le i\le n} \sum_{j=1}^{n} a_{ij}$$

и построим матрицу B такую, что

$$A \ge B \ge 0 \ \text{ и } \sum_{j=1}^{n} b_{ij} \equiv \alpha,\ \ i = \overline{1,n},$$

например, при $\alpha = 0$ полагаем $B = 0$, а если $\alpha > 0$, то можно взять

$$b_{ij} = \alpha\, a_{ij} \Bigl(\sum_{j=1}^{n} a_{ij}\Bigr)^{-1}.$$

Тогда из следствия 9.1 и леммы 9.1 вытекает

$$\rho(A) \ge \rho(B) = \alpha = \min_{1\le i\le n} \sum_{j=1}^{n} a_{ij}.$$

Аналогично доказывается верхняя оценка в первом неравенстве. Оценки для столбцовых сумм вытекают из оценок для строчных сумм матрицы A^T. □

Следствие 9.2. *Пусть* $A \in \mathbb{R}^{n\times n}$, $A \ge 0$ *и* $\sum\limits_{j=1}^{n} a_{ij} > 0$, $\forall i = \overline{1,n}$. *Тогда* $\rho(A) > 0$. *В частности,* $\rho(A) > 0$, *если* $A > 0$ *или если* A *неразложима и неотрицательна.*

Поскольку $\rho(S^{-1}AS) = \rho(A)$ для любой невырожденной матрицы S, доказанную выше теорему можно обобщить за счет введения некоторых свободных параметров.

Теорема 9.3. *Пусть* $A \in \mathbb{R}^{n\times n}$ *и предположим, что* $A \ge 0$. *Тогда для любого положительного вектора* $x \in \mathbb{C}^n$ *справедливы неравенства :*

$$\min_{1\le i\le n}\ \frac{1}{x_i} \sum_{j=1}^{n} a_{ij}\, x_j \le \rho(A) \le \max_{1\le i\le n}\ \frac{1}{x_i} \sum_{j=1}^{n} a_{ij}\, x_j,$$

$$\min_{1\le j\le n}\ x_j \sum_{j=1}^{n} \frac{a_{ij}}{x_i} \le \rho(A) \le \max_{1\le j\le n}\ x_j \sum_{j=1}^{n} \frac{a_{ij}}{x_i}.$$

Доказательство. Используем инвариантность спектрального радиуса относительно преобразования подобия, т.е., для любой невырожденной матрицы S справедливо:

$$\rho\left(S^{-1}AS\right) = \rho\left(A\right).$$

Возьмем

$$S = \mathrm{diag}(x_1, x_2, \ldots, x_n), \; x_i > 0, \; \forall i = \overline{1, n}.$$

Тогда при $A \geq 0$ верно $S^{-1}AS \geq 0$. Применяя к матрице $S^{-1}AS = [a_{ij}x_j x_i^{-1}]$ теорему 9.2, убеждаемся в справедливости теоремы 9.3. □

Упражнение 9.3. Доказать, что оценки теоремы 9.3 следуют из оценок следствия 8.1 стр. 270.

Следствие 9.3. *Пусть $A \in \mathbb{R}^{n\times n}$, $x \in \mathbb{R}^n$, и предположим, что $A \geq 0$ и $x > 0$. Если числа $\alpha, \beta \geq 0$ таковы, что*

$$\alpha x \leq Ax \leq \beta x,$$

то

$$\alpha \leq \rho(A) \leq \beta.$$

Если $\alpha x < Ax$, то $\alpha < \rho(A)$; если $Ax < \beta x$, то $\rho(A) < \beta$.

Доказательство. Если $\alpha x \leq Ax$, то

$$\alpha \leq \min_{1\leq i\leq n} x_i^{-1} \sum_{j=1}^{n} a_{ij}x_j$$

и по теореме 9.3 $\alpha \leq \rho(A)$. Если $\alpha x < Ax$, то для какого-то $\alpha' > \alpha$ имеем $\alpha' x \leq Ax$. В этом случае $\rho(A) \geq \alpha' > \alpha$, так что $\rho(A) > \alpha$. Верхние оценки проверяются аналогично. □

Следствие 9.4. *Пусть $A \in \mathbb{R}^{n\times n}$ и $A \geq 0$. Если A имеет положительный собственный вектор, то отвечающее ему собственное значение есть $\rho(A)$. Другими словами, если*

$$Ax = \lambda x, \; x > 0 \text{ и } A \geq 0,$$

то

$$\lambda = \rho(A).$$

Доказательство. Если $x > 0$ и $Ax = \lambda x$, то $\lambda \geq 0$ и $\lambda x \leq Ax \leq \lambda x$, но тогда согласно следствию 9.3, $\lambda \leq \rho(A) \leq \lambda$, что и доказывает следствие 9.4. □

9.2. Положительные и неотрицательные неразложимые матрицы

Лемма 9.2. *Пусть $A \in \mathbb{R}^{n\times n}$ и $A > 0$, $Ax = \lambda x$, $x \neq 0$ и $|\lambda| = \rho(A)$. Тогда $A\,|x| = \rho(A)\,|x|$ и $|x| > 0$.*

Доказательство. Очевидно, справедлива цепочка соотношений:

$$\rho(A)\,|x| = |\lambda|\,|x| = |\lambda x| = |Ax| \overset{2^0}{\leq} |A|\,|x| = A\,|x|\,,$$

так что $y \equiv A\,|x| - \rho(A)\,|x| \geq 0$. Поскольку $|x| \geq 0$ и $x \neq 0$, то по свойству 5^0 неотрицательных матриц справедливо $A\,|x| > 0$. Из следствия 9.3 тогда следует $\rho(A) > 0$. Поэтому если $y = 0$, то $A\,|x| = \rho(A)\,|x|$ и $|x| = \rho(A)^{-1}A\,|x| > 0$. Покажем, что $y \neq 0$ выполняться не может. Действительно, если $y \neq 0$, то положим $z \equiv A\,|x| > 0$ и, согласно 5^0, получаем

$$0 < Ay = Az - \rho(A)z, \; \text{или} \quad Az > \rho(A)z,$$

откуда, опираясь на следствие 9.3, выводим неравенство $\rho(A) > \rho(A)$, которое выполняться не может. Значит, $y = 0$, что и требовалось доказать. □

Теорема 9.4. *Пусть $A \in \mathbb{R}^{n\times n}$ и $A > 0$. Тогда $\rho(A) > 0$, $\rho(A)$ — собственное значение для A и существует положительный вектор x, такой, что $Ax = \rho(A)x$.*

Доказательство. Рассмотрим собственное значение λ, такое, что $|\lambda| = \rho(A) > 0$, и отвечающий ему собственный вектор $x \neq 0$. По лемме 9.2 искомым вектором будет $|x|$. □

Лемма 9.3. *Пусть $A \in \mathbb{R}^{n\times n}$ и $A > 0$, $Ax = \lambda x$, $x \neq 0$ и $|\lambda| = \rho(A)$. Тогда для некоторого $\theta \in \mathbb{R}$ имеем $e^{-i\theta}x = |x| > 0$.*

Доказательство приведено в [10, стр. 586, лемма 8.2.3.]. □

Теорема 9.5. *Пусть $A \in \mathbb{R}^{n\times n}$ и $A > 0$. Тогда любое собственное значение $\lambda \neq \rho(A)$ удовлетворяет неравенству $|\lambda| < \rho(A)$.*

Доказательство. По определению $|\lambda| \leq \rho(A)$ для всех $\lambda \in \sigma(A)$. Пусть $|\lambda| = \rho(A)$ и $Ax = \lambda x$, $x \neq 0$. Согласно лемме 9.3, $Aw = \lambda w$, где $w = e^{-i\theta}x > 0$ для какого-то $\theta \in R$. Отсюда, опираясь на следствие 9.4, получаем $\lambda = \rho(A)$. □

Теорема 9.6. *Если $A \in \mathbb{R}^{n\times n}$ и $A > 0$, то $\rho(A)$ есть собственное значение алгебраической кратности 1; другими словами, $\rho(A)$ — это простой корень характеристического уравнения $p_A(t) = 0$, где $p_A(t)$ — характеристический полином матрицы A.*

Доказательство. Согласно теореме Шура о триангуляризации 5.5 стр. 146 матрицу A можно представить в виде $A = UTU^*$, где U — унитарная, а T — верхнетреугольная матрица с диагональными элементами $\rho, \ldots, \rho, \lambda_{k+1}, \ldots, \lambda_n$, где $\rho = \rho(A)$ — собственное значение алгебраической кратности $k \geq 1$; при $i = k+1, \ldots, n$ собственное значение λ_i по модулю строго меньше $\rho(A)$. В этом случае получаем

$$L = \lim_{m\to\infty} [\rho(A)^{-1}A]^m =$$

$$= U \lim_{m\to\infty} \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & \frac{\lambda_{k+1}}{\rho} & & \\ & & & & \ddots & \\ & & & & & \frac{\lambda_n}{\rho} \end{bmatrix}^m U^* =$$

$$= U \begin{bmatrix} 1 & & & & & * \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ 0 & & & & & 0 \end{bmatrix} U^*,$$

где в последних двух выражениях диагональный элемент 1 повторяется k раз, а элемент 0 на диагонали в последнем выражении повторяется $n - k$ раз. Верхняя треугольная матрица в последнем выражении имеет ранг не меньше k, и, так как для L ранг равен 1, случай $k > 1$ невозможен. □

Теорема 9.7 (Перрона). *Если $A \in \mathbb{R}^{n\times n}$ и $A > 0$, то*

a) $\rho(A) > 0$*;*

b) $\rho(A)$ *есть собственное значение матрицы* A $(\rho(A) = \lambda_A \in \sigma(A))$*;*

c) для некоторого $x_A \in \mathbb{C}^n$ *имеем* $x_A > 0$ *и* $Ax_A = \rho(A)x_A$*;*

d) $\rho(A)$ *есть алгебраически (а значит, и геометрически) простое собственное значение для* A*;*

e) $|\lambda| < \rho(A)$ *для всякого собственного значения* $\lambda \neq \rho(A)$*; другими словами, только одно собственное значение* λ_A*, равное именно* $\rho(A)$*, имеет максимальный модуль;*

f) $\lim\limits_{m\to\infty} \left[\rho(A)^{-1}A\right]^m = L > 0$, *где*

$$L \equiv xy^T,\ Ax = \rho(A)x,\ A^Ty = \rho(A)y,\ x > 0,\ y > 0,\ x^Ty = 1.$$

Доказательство. Положения a)—e) следуют из теорем 9.4—9.6. Доказательство положения f) можно найти в [10, стр. 589, теорема 8.2.8]. □

Определение 9.2. Единственный нормированный $\left(\sum\limits_{i=1}^{n} x_i = 1\right)$ собственный вектор x из утверждения c) теоремы Перрона называется *перроновым вектором* матрицы A, число $\rho(A)$ — *перроновым корнем* матрицы A. Перронов вектор y для матрицы A^T называется *левым перроновым вектором* матрицы A.

Пример 9.2 (Перронов корень и перронов вектор). Найдем $\lim\limits_{m\to\infty} [\rho(A)^{-1}A]^m$ для матрицы

$$A = \begin{bmatrix} 2 & 2 \\ 1 & 3 \end{bmatrix} > 0.$$

Матрица A имеет два собственных значения: $\lambda_1 = 1$ и $\lambda_2 = 4$. Перронов корень матрицы A — это число $\lambda_A = \rho(A) = 4$, перронов вектор —

$$x = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}.$$

Левый собственный вектор матрицы A, соответствующий перронову корню, найдем из равенства $y^TA = \lambda_A y^T$:

$$y = \begin{bmatrix} \alpha \\ 2\alpha \end{bmatrix}.$$

Подберем параметр α из условия нормировки f) теоремы Перрона:

$$x^Ty = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} \alpha \\ 2\alpha \end{bmatrix} = \frac{3}{2}\alpha = 1,$$

откуда $\alpha = \frac{2}{3}$ и

$$y = \begin{bmatrix} \frac{2}{3} \\ \frac{4}{3} \end{bmatrix}.$$

Тогда матрица L из теоремы Перрона имеет вид

$$L = xy^T = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix} \begin{bmatrix} \frac{2}{3} & \frac{4}{3} \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} \end{bmatrix}.$$

Теорема 9.8 (Фань Цзы). *Пусть $A \in \mathbb{C}^{n\times n}$ и $B \in \mathbb{R}^{n\times n}$, $B \geq 0$ и $B \geq |A|$. Тогда любое собственное значение матрицы A принадлежит области*

$$\bigcup_{i=1}^{n} \Big\{ z \in \mathbb{C} : |z - a_{ii}| \leq \rho(B) - b_{ii} \Big\}. \tag{9.1}$$

Доказательство. Будем считать, что $B > 0$. Если это не так, то можно рассматривать матрицу $B_\varepsilon \equiv [b_{ij} + \varepsilon]$, где $\varepsilon > 0$. При этом $B_\varepsilon > |A|$ и $\rho(B_\varepsilon) - (b_{ii} + \varepsilon) \underset{\varepsilon\to 0}{\to} \rho(B) - b_{ii}$.

По теореме Перрона для какого-то положительного вектора x имеем $Bx = \rho(B)x$. Рассматривая это равенство покомпонентно для $\forall\, i = \overline{i, n}$, имеем

$$\rho(B)x_i = \sum_{j\neq i} b_{ij}x_j + b_{ii}x_i \geq \sum_{j\neq i} |a_{ij}|\, x_j + b_{ii}x_i \quad \overset{x>0}{\Rightarrow}$$

$$\Rightarrow \quad \frac{1}{x_i} \sum_{j\neq i} |a_{ij}|\, x_j \leq \rho(B) - b_{ii},\ i = \overline{1, n}.$$

Положив в (8.4) $p_i = x_i$, убеждаемся в (9.1). □

Пример 9.3 (Область Фань Цзы). Рассмотрим матрицы

$$A = \begin{bmatrix} -1 & 0 \\ 3 & 2 \end{bmatrix},\ B = \begin{bmatrix} 1 & 0 \\ 3 & 2 \end{bmatrix}.$$

Матрица A имеет два собственных значения: $\lambda_1 = -1$, $\lambda_2 = 2$, спектральный радиус матрицы B равен $\rho(B) = 2$. Область (9.1) образуется двумя кругами: с центром в точке $a_{11} = -1$ и радиусом $\rho(B) - b_{11} = 2 - 1 = 1$, а также с центром в точке $a_{22} = 2$ и радиусом $\rho(B) - b_{22} = 2 - 2 = 0$. Она изображена на рис.9.1, собственные значения матрицы A указаны на нем точками.

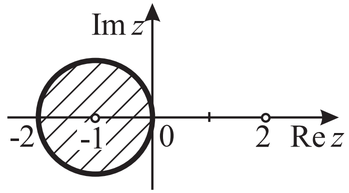


Рис. 9.1

Для сравнения строчная и столбцовая области Гершгорина, содержащие все собственные значения матрицы A, изображены на рис. 9.2.

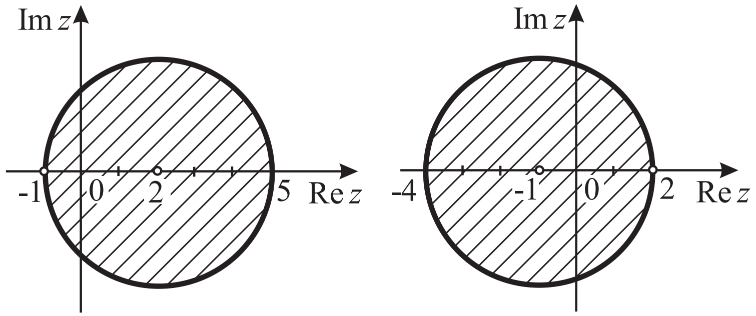


Рис. 9.2

Следующая теорема является обобщением теоремы Перрона на неотрицательные неразложимые матрицы.

Теорема 9.9 (Перрона— Фробениуса). *Пусть $A \in \mathbb{R}^{n\times n}$ неразложима и $A \geq 0$. Тогда*

a) $\rho(A) > 0$;

b) $\rho(A)$ есть собственное значение матрицы A:

$$\rho(A) = \lambda_A \in \sigma(A);$$

c) для некоторого $x_A \in \mathbb{C}^n$ имеем $x_A > 0$ и $Ax_A = \rho(A)x_A$;

d) $\rho(A)$ есть алгебраически (а значит, и геометрически) простое собственное значение для A.

Определение 9.3. Число $\lambda_A = \rho(A)$ называют *числом Фробениуса*, а соответствующий собственный вектор x_A — *вектором Фробениуса.*

Сопоставим всякой неразложимой матрице $A \geq 0$ матрицу $\tilde{A} = \lambda_A^{-1}A$. Матрица $\tilde{A}$ неразложима и $\lambda_{\tilde{A}} = 1$, а вектор Фробениуса для нее совпадает с x_A.

Определение 9.4. Неразложимую неотрицательную матрицу A называют *устойчивой,* если для любого вектора $x \in \mathbb{R}^n$, $x \neq 0$, последовательность $\tilde{A}^k x,\ k = 1, 2, \ldots$, сходится.

Теорема 9.10. *Если $\tilde{A}^k x,\ k = 1, 2, \ldots$, сходится, то*

$$\lim_{k\to\infty} \tilde{A}^k x = \frac{\|x\|}{\|x_A\|} x_A,$$

где норма векторов x введена следующим образом:

$$\|x\| = y_A^T\,|x|\,.$$

Здесь y_A — вектор Фробениуса для матрицы A^T, $|x|$ — вектор, составленный из абсолютных значений компонент вектора x.

Упражнение 9.4. Используя определение векторной нормы и положительность вектора Фробениуса, доказать, что норма, введенная в теореме 9.10, действительно является нормой.

Теорема 9.11. *Неразложимая неотрицательная матрица A устойчива тогда и только тогда, когда $|\lambda| < \lambda_A$ для любого ее собственного значения, отличного от λ_A.*

Пример 9.4 (Устойчивая матрица). Неразложимая неотрицательная матрица

$$A=\begin{bmatrix} 1 & 3 \\ 2 & 0 \end{bmatrix}$$

имеет собственные значения

$$\lambda_1(A)=3,\ \lambda_2(A)=-2,$$

число Фробениуса

$$\lambda_A = \rho(A) = 3$$

и вектор Фробениуса

$$x_A = \begin{bmatrix} 3 \\ 2 \end{bmatrix}.$$

Матрица

$$\tilde{A} = \begin{bmatrix} \frac{1}{3} & 1 \\ \frac{2}{3} & 0 \end{bmatrix}$$

неразложима, имеет число Фробениуса

$$\lambda_{\tilde{A}} = \rho(\tilde{A}) = 1$$

и вектор Фробениуса

$$x_{\tilde{A}} = x_A.$$

Так как

$$|\lambda_2(A)| = 2 < \lambda_A = \lambda_1(A) = 3,$$

то, согласно теореме 9.11, матрица A устойчива.

Найдем

$$\lim_{k\to\infty} \tilde{A}^k x = \frac{\|x\|}{\|x_A\|} x_A$$

для $x = [4,\, 6]^T$. Вектор Фробениуса y_A для матрицы A^T равен

$$y_A = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Тогда

$$\|x\| = [1,\, 1] \begin{bmatrix} 4 \\ 6 \end{bmatrix} = 10, \quad \|x_A\| = [1,\, 1] \begin{bmatrix} 3 \\ 2 \end{bmatrix} = 5$$

и

$$\lim_{k\to\infty} \tilde{A}^k x = \frac{10}{5} \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 6 \\ 4 \end{bmatrix}.$$

Заметим, что непосредственные вычисления также дают

$$\lim_{k\to\infty} \tilde{A}^k x = \begin{bmatrix} 0,6 & 0,6 \\ 0,4 & 0,4 \end{bmatrix} \begin{bmatrix} 4 \\ 6 \end{bmatrix} = \begin{bmatrix} 6 \\ 4 \end{bmatrix}.$$

9.3. Примитивные матрицы

Определение 9.5. Пусть неразложимая квадратная матрица $A \geq 0$ имеет k собственных значений, равных по модулю $\rho(A)$. Тогда если $k = 1$, то матрица A называется *примитивной*. В противном случае A называется *импримитивной* с *индексом импримитивности* k.

Пример 9.5 (Примитивная и импримитивная матрицы).

Матрица

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

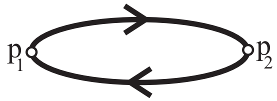


Рис. 9.3

неразложима, так как ее граф, изображенный на рис. 9.3, является сильносвязным. A имеет два собственных значения: $\lambda_1 = -1$, $\lambda_2 = 1$. Так как $|\lambda_1| = |\lambda_2| = 1$, то индекс импримитивности матрицы A равен $k = 2$, а значит, матрица A импримитивная.

Матрица

$$B = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$$

неразложима, так как ее граф, изображенный на рис. 9.4,

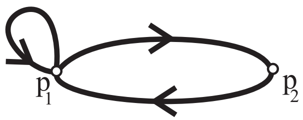


Рис. 9.4

сильносвязен. B имеет два собственных значения:

$$\lambda_1 = \frac{1-\sqrt{5}}{2},\ \lambda_2 = \frac{1+\sqrt{5}}{2}.$$

Так как $|\lambda_1| \neq |\lambda_2|$, то индекс импримитивности матрицы B равен $k = 1$, а значит, матрица A примитивная.

Теорема 9.12. *[10, стр. 608]. Если $A \in \mathbb{R}^{n\times n}$ неотрицательна и примитивна, то*

$$\lim_{m\to\infty} \left[\rho(A)^{-1}A\right]^{\,m} = L > 0,$$

где $L = xy^T$, $Ax = \rho(A)x$, $A^Ty = \rho(A)y$, $x > 0$, $y > 0$, $x^Ty = 1$.

Таким образом, все пункты теоремы Перрона, сформулированной для класса положительных матриц, обобщены на класс примитивных неотрицательных матриц.

Упражнение 9.5. Для неотрицательных примитивных матриц сравнить пределы, получаемые согласно теореме 9.10 и с применением теоремы 9.12. Доказать, что эти пределы совпадают.

Для практической проверки примитивности заданной неотрицательной матрицы удобно уметь делать это без явного вычисления собственных значений. Далее дано несколько таких полезных результатов.

Вычислить индекс импримитивности матрицы A можно следующим образом.

Теорема 9.13. *[3, стр. 252] Пусть*

$$p_A(\lambda) = \lambda^n + a_1\lambda^{n_1} + \ldots + a_t\lambda^{n_t} \qquad (9.2)$$

— характеристический полином матрицы A, причем $a_1, a_2, \ldots, a_t$ ненулевые и $n > n_1 > \ldots > n_t \geq 0$. Вычисляем разности

$$n - n_1, n_1 - n_2, \ldots, n_{t-1} - n_t.$$

Тогда индекс импримитивности k матрицы A равен наибольшему общему делителю разностей $n - n_j$, $j = \overline{1,t}$.

Пример 9.6 (Нахождение индекса импримитивности). Если

$$p(\lambda) = \lambda^{10} + a_1\lambda^7 + a_2\lambda, \quad a_1,\ a_2 \neq 0,$$

то

$$n - n_1 = 10 - 7 = 3,\ n - n_2 = 10 - 1 = 9$$

и
$$k = 3.$$

Если
$$p(\lambda) = \lambda^{10} + a_1\lambda^7 + a_2\lambda + a_3, \quad a_1,\ a_2,\ a_3 \neq 0,$$
то
$$n - n_1 = 10 - 7 = 3,\ n - n_2 = 10 - 1 = 9,\ n - n_3 = 10 - 0 = 10$$
и
$$k = 1.$$

Пример 9.7 (Индекс примитивности). Для матрицы A из примера 9.5 имеем:
$$p_A(\lambda) = \lambda^2 - 1,\ n - n_1 = 2 - 0 = 2,\ k = 2,$$
т.е. матрица A импримитивна. Для матрицы B имеем:
$$p_B(\lambda) = \lambda^2 - \lambda - 1,\ n - n_1 = 2 - 1 = 1,\ n - n_2 = 2 - 0 = 2,\ k = 1,$$
т.е. матрица B примитивна.

Теорема 9.14. *Матрица $A \geq 0$ примитивна в том и только в том случае, когда $A^m > 0$ для некоторого $m \geq 1$.*

Доказательство. Предположим, что $A^m > 0$, и, следовательно, A^m неразложима (см. следствие 8.2 стр. 286). Далее, из разложимости A следовала бы разложимость A^m (см. упражнение 8.5 стр. 283), а следовательно, $A^m > 0$ означает, что A неразложима. Если бы A имела индекс импримитивности $k > 1$, то так как собственные значения для A^m являются m-ми степенями собственных значений для A, A^m также имела бы индекс импримитивности $k > 1$. Но это противоречило бы теореме

Перрона, примененной к положительной матрице A^m. Следовательно, $k = 1$ и A примитивна.

Обратно, если A примитивна, то $\lim\limits_{m\to\infty} \left[\rho(A)^{-1}A\right]^{m} = L > 0$, согласно теореме 9.12, и так как $\rho(A) > 0$, то для какого-то $m \geq 1$ верно $A^m > 0$. □

Определение 9.6. Для любой заданной примитивной матрицы A наименьшее число k такое, что $A^k > 0$, называется ее *индексом примитивности.*

Пример 9.8 (Проверка примитивности). Для матрицы A из примера 9.5 имеем:
$$A^{2l} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = E_2,\ A^{2l+1} = A,\ l = 0, 1, 2, \ldots.$$
Значит, для любого $m \geq 1$ матрица A^m не является положительной и согласно теореме 9.14 матрица A импримитивна. Для матрицы B из примера 9.5 при $m = 2$ получаем
$$B^2 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} > 0$$
и согласно теореме 9.14 матрица B примитивна.

Пусть $\Gamma(A)$ — орграф неразложимой матрицы $A \geq 0$, S_i — множество всех замкнутых путей этого графа, содержащих вершину i, l_i — множество длин всех путей из S_i, g_i — наибольший общий делитель всех длин из l_i.

Теорема 9.15 (Графовый критерий примитивности). *[10, стр. 609] Неразложимая матрица $A \geq 0$ примитивна тогда и только тогда, когда $g_i = 1$, $\forall i = \overline{1, n}$.*

Пример 9.9 (Проверка примитивности по графовому критерию). Для матрицы A из примера 9.5 и ее орграфа имеем:

$$S_1 = \{1 \to 2 \to 1\}\,,\ S_2 = \{2 \to 1 \to 2\}\,,$$

$$l_1 = \{2\}\,,\ l_2 = \{2\}\,,$$

$$g_1 = g_2 = 2 \neq 1,$$

следовательно, A импримитивна.

Для матрицы B из примера 9.5 и ее орграфа имеем:

$$S_1 = \{1 \to 1,\ 1 \to 2 \to 1\}\,,\ S_2 = \{2 \to 1 \to 2,\ 2 \to 1 \to 1 \to 2\}\,,$$

$$l_1 = \{1, 2\},\ l_2 = \{2, 3\}\,,\quad g_1 = g_2 = 1,$$

следовательно, B примитивна.

Теорема 9.16. *[10, стр. 612],[9, стр. 356] Если матрица $A \in \mathbb{R}^{n\times n}$ неотрицательна, то A примитивна тогда и только тогда, когда $A^{n^2-2n+2} > 0$.*

Пример 9.10 (Проверка примитивности). Матрица

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \in \mathbb{R}^{2\times 2}$$

имеет своим графом орграф $\Gamma(A)$, изображенный на рисунке 9.5.

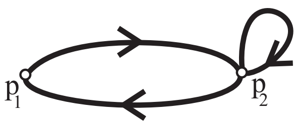


Рис. 9.5

Граф $\Gamma(A)$ сильносвязный, следовательно, матрица A неразложима. Вычислим

$$A^{2^2-2\times2+2} = A^2 = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} > 0.$$

Значит, согласно теореме 9.16, матрица A примитивна. Непосредственное вычисление дает собственные значения:

$$\lambda_1 = \frac{1-\sqrt{5}}{2},\ \lambda_2 = \frac{1+\sqrt{5}}{2},\ |\lambda_1| < |\lambda_2|\,,$$

значит индекс импримитивности $k{=}1$ и матрица A примитивна.

Свойства примитивных матриц

1^0. Если $A \ge 0$ — примитивная матрица, то A^m является неотрицательной, неразложимой и примитивной для всех $m = 1, 2, \ldots$.

2^0. Если $A \ge 0$ — неразложимая матрица с индексом импримитивности k, то существует матрица перестановок P, такая, что
$$PA^kP^T = \sum_{i=1}^{k} A_i,$$
где A_i — примитивные матрицы с одним и тем же максимальным собственным значением.

3^0. Если $A > 0$, то она примитивна.

4^0. Если $A \ge 0$ неразложима и хотя бы один ее диагональный элемент положителен, то A примитивна.

5^0. Если A — произвольная неразложимая матрица, B — любая неотрицательная матрица с положительным следом, то их сумма $A + B$ является примитивной матрицей.

6^0. Если $A \in \mathbb{R}^{n\times n}$ неотрицательна и неразложима и имеет d положительных элементов на главной диагонали, где $1 \le d \le n$, то индекс примитивности $k \le 2n - d - 1$.

7^0. Примитивная матрица всегда устойчива.

Упражнение 9.6. Доказать свойства примитивных матриц.

9.4. Стохастические и двоякостохастические матрицы

Определение 9.7. Неотрицательная матрица $A \in \mathbb{R}^{n\times n}$, для которой все строчные суммы равны $+1$, называется *(строчной) стохастической матрицей. Столбцовая стохастическая матрица* — это матрица, транспонированная к строчной стохастической матрице.

Стохастическая матрица $A \in \mathbb{R}^{n\times n}$, для которой A^T также является стохастической, называется *двоякостохастической.*

Пример 9.11 (Стохастическая матрица). Матрица

$$A = \begin{bmatrix} 0,5 & 0,3 & 0,2 \\ 0,7 & 0 & 0,3 \\ 0 & 0,5 & 0,5 \end{bmatrix}$$

является стохастической, но не является двоякостохастической.

Матрица

$$B = \begin{bmatrix} 0,3 & 0,5 & 0 & 0,2 \\ 0,7 & 0 & 0,3 & 0 \\ 0 & 0,2 & 0 & 0,8 \\ 0 & 0,3 & 0,7 & 0 \end{bmatrix}$$

является двоякостохатической.

Пример 9.12 (Стохастичность матрицы перестановок). Любая матрица перестановок P является двоякостохастической. Например,

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Теорема 9.17. *Неотрицательная матрица A является стохастической тогда и только тогда, когда она имеет собственное значение 1 с правым собственным вектором*

$$e = \begin{bmatrix} 1 \\ 1 \\ \dots \\ 1 \end{bmatrix}.$$

Кроме того, спектральный радиус стохастической матрицы равен 1.

Упражнение 9.7. Провести доказательство теоремы 9.17. Использовать лемму 9.1.

Широкий класс неотрицательных матриц может быть при помощи преобразования подобия сведен к стохастическим матрицам.

Теорема 9.18*. *Пусть A — неотрицательная матрица с максимальным собственным значением λ. Пусть существует положительный правый собственный вектор x, соответствующий λ. Положим $X = \operatorname{diag}\{x_1, \dots, x_n\}$. Тогда*

$$A = \lambda X P X^{-1},$$

где P — стохастическая матрица.

Доказательство. Пусть $P = \lambda^{-1}X^{-1}AX$. Покажем, что P — стохастическая. Так как по определению собственного вектора и собственного значения $Ax = \lambda x$, то

$$\sum_{j=1}^{n} a_{ij}x_j = \lambda x_i, \;\; i = \overline{1, n}.$$

По определению $p_{ij} = \lambda^{-1}x_i^{-1}a_{ij}x_i$, и, следовательно $\sum\limits_{j=1}^{n} p_{ij} = \lambda^{-1}x_i^{-1}\sum\limits_{j=1}^{n} a_{ij}x_i = 1$, а значит, P — стохастическая. □

Пример 9.13 (Разложение стохастической матрицы). Матрица

$$A = \begin{bmatrix} 1 & 0 & 3 \\ 0 & 2 & 1 \\ 4 & 0 & 1 \end{bmatrix} \geq 0$$

имеет максимальное собственное значение $\lambda = 1 + 2\sqrt{3}$ и соответствующий положительный собственный вектор $x = \left[\frac{\sqrt{3}}{2}, \frac{1+2\sqrt{3}}{11}, 1\right]^T$.

Матрицы из теоремы 9.18 имеют вид:

$$X = \operatorname{diag}\left\{\frac{\sqrt{3}}{2}, \frac{1+2\sqrt{3}}{11}, 1\right\}, \; X^{-1} = \operatorname{diag}\left\{\frac{2}{\sqrt{3}}, \frac{11}{1+2\sqrt{3}}, 1\right\},$$

$$P = \frac{1}{\lambda}X^{-1}AX = \begin{bmatrix} \frac{1}{1+2\sqrt{3}} & 0 & \frac{2\sqrt{3}}{1+2\sqrt{3}} \\ 0 & \frac{2}{1+2\sqrt{3}} & \frac{11}{\left[1+2\sqrt{3}\right]^2} \\ \frac{2\sqrt{3}}{1+2\sqrt{3}} & 0 & \frac{1}{1+2\sqrt{3}} \end{bmatrix}.$$

Легко проверить, что суммы элементов в строках матрицы P равны 1, т.е. матрица P является стохастической.

Теорема 9.19 (Биркгофа). *Матрица $A \in \mathbb{R}^{n\times n}$ является двоякостохастической в том и только том случае, когда для некоторого $N < \infty$ существуют матрицы перестановок $P_1, \ldots, P_N \in \mathbb{R}^{n\times n}$ и положительные числа $\alpha_1, \ldots, \alpha_N \in \mathbb{R}$, такие, что $\alpha_1 + \ldots + \alpha_N = 1$ и $A = \alpha_1 P_1 + \ldots + \alpha_N P_N$.*

Доказательство.* Достаточность очевидна. Докажем необходимость. Пусть $A \in \mathbb{R}^{n\times n}$ — двоякостохастическая. Возможны два случая: 1) A — матрица перестановок, 2) не является матрицей перестановок.

В первом случае покажем, что любая матрица перестановок является крайней точкой множества двоякостохастических матриц. Действительно, если бы можно было записать $a = \alpha_1 B + \alpha_2 C$, где $0 < \alpha_1, \alpha_2 < 1$, $\alpha_1 + \alpha_2 = 1$ и B, C — двоякостохастические матрицы, то для элементов в B и C, отвечающих элементу $a_{ij} = 0$, было бы верно $0 = a_{ij} = \alpha_1 b_{ij} + \alpha_2 c_{ij}$, откуда $b_{ij} = c_{ij} = 0$, так как $\alpha_1 \neq 0, \alpha_2 \neq 0$, $b_{ij} \geq 0$, $c_{ij} \geq 0$. Поскольку матрицы B и C двоякостохастические, то их ненулевые элементы должны равняться $+1$ и занимать те же позиции, что ненулевые элементы в A. Итак, $A = B = C$. Это доказывает, что любая матрица перестановки является крайней точкой множества двоякостохастических матриц.

Рассмотрим случай 2). Тогда хотя бы в одной строке (например, i_1), есть по меньшей мере два ненулевых элемента. Возьмем ненулевой элемент $a_{i_1 i_2}$. Так как сумма элементов строки i_1 равна $+1$, все элементы неотрицательны и в строке i_1 не менее двух ненулевых элемента, то $0 < a_{i_1 i_2} < 1$. Далее, так как сумма элементов столбца i_2 равна $+1$, то в этом столбце должен найтись еще один ненулевой элемент $a_{i_3 i_2}$, $i_3 \neq i_2$, и для него $0 < a_{i_3 i_2} < 1$. По той же причине в одной строке с элементом $a_{i_3 i_2}$ имеется другой ненулевой элемент $a_{i_3 i_4}$, $i_4 \neq i_3$, и для него $0 < a_{i_3 i_4} < 1$.

Продолжая этот процесс, будем помечать выбранные элементы. После конечного числа шагов обязательно возникнет ситуация, когда мы выберем элемент, который ранее уже помечался. Последовательность элементов от первого до второго появления элемента a_{ij} (включая первое, но не второе появление) — это конечная упорядоченная последовательность элементов матрицы A, в которой любая пара соседних элементов находится попеременно то в одном столбце, то в одной строке. Обозначим $a_{i'j'}$ наименьший (положительный) элемент в этой последовательности. Построим матрицу $B \in \mathbb{R}^{n\times n}$, в которой в позиции, соответствующей первому элементу a_{ij} рассмотренной последовательности, поставим $+1$, в позиции второго элемента поставим -1, далее $+1$ и так далее поочередно. Все остальные элементы положим равными нулю. Заметим, что все столбцовые и строчные суммы матрицы B равны нулю.

Положим $A_+ = A + \frac{1}{2}a_{i'j'}B$, $A_- = A - \frac{1}{2}a_{i'j'}B$. Заметим, что $A_+ \geq 0$, $A_- \geq 0$ (в силу минимальности элемента $a_{i'j'}$) и их строчные и столбцовые суммы равны $+1$ (так как для B они равны 0). Поэтому матрицы A_+, A_- являются двоякостохастическими. Тогда $A = \frac{1}{2}A_+ + \frac{1}{2}A_-$ и $A_+ \neq A_-$, а это означает, что матрица A не является крайней точкой для множества двоякостохастических матриц.

Проведенное рассуждение показывает, что матрица является крайней точкой компактного выпуклого множества двоякостохастических матриц в том и только том случае, когда она есть матрица перестановки. Утверждение теоремы вытекает из того факта, что любая точка произвольного выпуклого компактного множества есть выпуклая комбинация его крайних точек. □

Пример 9.14 (Разложение двоякостохастической матрицы). Матрица B из примера 9.11 представима в виде

$$A = \alpha_1 P_1 + \alpha_2 P_2 + \alpha_3 P_3,$$

где $\alpha_1 = 0,3$, $\alpha_2 = 0,5$, $\alpha_3 = 0,2$,

$$P_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \ P_2 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \ P_3 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

Теорема 9.20. *Если $A \in \mathbb{R}^{n\times n}$ — неразложимая стохастическая матрица, то матрица $A^\infty = \lim\limits_{m\to\infty} A^m$ существует тогда и только тогда, когда A примитивна.*

Доказательство следует из теорем 9.12 и 9.17. □

Пример 9.15 (Предел степеней стохастической матрицы). Проверим, существует ли $A^\infty = \lim\limits_{m\to\infty} A^m$ для стохастической матрицы A из примера 9.11. Орграф $\Gamma(A)$ матрицы A изображен на рисунке 9.6.

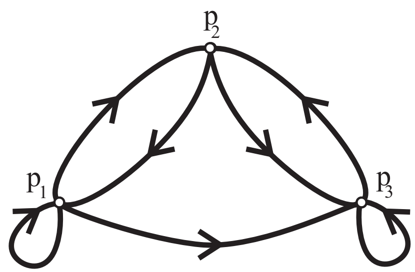


Рис. 9.6

Так как он является сильносвязным, то матрица A неразложима. Далее для проверки примитивности рассматриваемой матрицы воспользуемся теоремой 9.16. Находим

$S_1 = \{1 \to 1,\ 1 \to 2 \to 1\}$, $l_1 = \{1, 2\}$, $g_1 = 1$,

$S_2 = \{2 \to 1 \to 3 \to 2,\ 2 \to 3 \to 2\}$, $l_2 = \{3, 2\}$, $g_2 = 1$,

$S_3 = \{3 \to 3,\ 3 \to 2 \to 3,\ 3 \to 2 \to 1 \to 3\}$, $l_3 = \{1, 2, 3\}$, $g_3 = 1$, откуда следует, что матрица A примитивна, а значит, согласно теореме 9.20, существует $A^\infty = \lim\limits_{m\to\infty} A^m$.

Для вычисления A^∞ воспользуемся теоремой 9.12.

Во-первых, согласно теореме 9.17 число Фробениуса матрицы A равно $\rho(A) = 1$, а соответствующий вектор Фробениуса имеет вид

$$x = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

Далее находим левый собственный вектор, соответствующий числу Фробениуса и нормированный условием $x^T y = \sum y_i = 1$:

$$y^T = \begin{bmatrix} \frac{35}{89} & \frac{25}{89} & \frac{29}{89} \end{bmatrix}.$$

Тогда

$$A^\infty = \lim_{m\to\infty} A^m = xy^T =$$

$$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} \frac{35}{89} & \frac{25}{89} & \frac{29}{89} \end{bmatrix} = \begin{bmatrix} \frac{35}{89} & \frac{25}{89} & \frac{29}{89} \\ \frac{35}{89} & \frac{25}{89} & \frac{29}{89} \\ \frac{35}{89} & \frac{25}{89} & \frac{29}{89} \end{bmatrix}.$$

Задачи и упражнения

1. Доказать свойства (a)—(j) неотрицательных матриц:
   (a) $|aA| = |a|\,|A|$ $\forall a \in \mathbb{C}$.
   (b) $|A + B| \le |A| + |B|$.
   (c) Если $A \ge 0$ и $A \ne 0$, то это ещё не означает, что $A > 0$, если хотя бы одно из чисел n или m больше 1.
   (d) Если $A \ge 0$, $B \ge 0$ и $a, b \ge 0$, то $aA + bB \ge 0$.
   (e) Если $A \ge B$, $C \ge D$, то $A + C \ge B + D$.
   (f) Если $A \ge B$, $B \ge C$, то $A \ge C$.
   (g) $|AB| \le |A|\,|B|$.
   (h) Если $0 \le A \le B$ и $0 \le C \le D$, то $0 \le AC \le BD$.

   (i) Если $A \ge 0$, то $A^m \ge 0$; если $A > 0$, то $A^m > 0$, $\forall\, m = 1, 2, \ldots$.
   (j) Если $A \ge 0$, $x > 0$ и $Ax = 0$, то $A = 0$.
2. Доказать следующее утверждение.

   Пусть $A \in \mathbb{R}^{n\times n}$ и $A \ge 0$. Если $\tilde{A}$ — произвольная главная подматрица (см. 6.4) в A, то $\rho(\tilde{A}) \le \rho(A)$. В частности, $\max\limits_{i=\overline{1,n}} a_{ii} \le \rho(A)$. Т.е., спектральный радиус любой главной подматрицы матрицы A не превосходит ее спектрального радиуса.
3. Найти спектральный радиус матрицы
$$A = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$
   и спектральные радиусы всех ее главных подматриц. Продемонстрировать выполнение утверждения из предыдущего задания.
4. Показать, что если $A, B \in \mathbb{R}^{n\times n}$ и $0 \le A < B$, то $\rho(A) < \rho(B)$.
5. Доказать, что если $A \ge 0, B \ge C$ и AB определено, то $AB \ge AC$.
6. Получить оценку спектрального радиуса матрицы
$$A = \begin{bmatrix} 3 & 0 & 1 & 0 \\ 1 & 6 & 1 & 1 \\ 2 & 0 & 2 & 0 \\ 1 & 3 & 1 & 4 \end{bmatrix},$$
   основываясь на теореме 9.2.
7. Доказать, что если $A \ge 0$ и $A^k > 0$ для некоторого k, то $\rho(A) > 0$.
8. Привести пример 2×2 матрицы A, такой, что $A \ge 0$ и $A^2 > 0$, хотя A не является положительной.
9. Построить матрицу B, которая подобна матрице
$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

и не имеет нулевых элементов. Каков ее спектральный радиус? Будет ли она неотрицательной? Как это связано с последней частью следствия 9.2?

10. Пусть $A \in \mathbb{R}^{n\times n}$ и $A > 0$. Доказать, что если для какого-то вектора $x \in \mathbb{C}^n$ справедливы соотношения $x \geq 0$, $x \neq 0$ и $Ax = \lambda x$, то x есть кратное перронова вектора для A и $\lambda = \rho(A)$.

11. Пусть $A > 0$, x — перронов вектор матрицы A и y — перронов вектор матрицы A^T. Доказать, что $x^T y > 0$.

12. Доказать следующее утверждение. Пусть $A \in \mathbb{R}^{n\times n}$ и $A \geq 0$. Если A имеет положительный собственный вектор, то
$$\max_{x>0} \min_{1\leq i\leq n} \frac{1}{x_i} \sum_{j=1}^{n} a_{ij}x_j \leq \rho(A) \leq \min_{x>0} \max_{1\leq i\leq n} \frac{1}{x_i} \sum_{j=1}^{n} a_{ij}x_j.$$

13. В упражнении 6 стр. 288 рассматривается некоторая положительная 2×2-матрица. Обсудить этот пример в свете упражнения 12.

14. Доказать, что если $A \in \mathbb{R}^{n\times n}$ и $A > 0$, то для нее справедливы неравенства из упражнения 12.

15. Доказать, что если $A > 0$ и x — перронов вектор для A, то $\rho(A) = \sum\limits_{i,j=1}^{n} a_{ij}x_j$.

16. Показать, что если положительная матрица невырожденная, то обратная к ней матрица не может быть неотрицательной.

17. Показать на примерах, что утверждения теоремы Перрона, за исключением b) и c), вошедших в теорему 9.3, в общем случае не верны для произвольных неотрицательных матриц.

18. Доказать, что если $A \geq 0$ и $A^k > 0$ для какого-то $k \geq 1$, то A имеет положительный собственный вектор.

19. Доказать, что если неотрицательная квадратная матрица с положительными диагональными элементами возводится в степень, то любой диагональный элемент остается таковым для всех более высоких степеней.

20. Выяснить, является ли заданная матрица примитивной:
a) $\begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 3 & 0 & 4 \\ 2 & 0 & 2 & 0 \end{bmatrix}$, b) $\begin{bmatrix} 2 & 1 & 1 & 3 \\ 0 & 4 & 0 & 1 \\ 0 & 1 & 0 & 4 \\ 0 & 2 & 0 & 3 \end{bmatrix}$,
c) $\begin{bmatrix} 0 & a & 0 & 0 \\ 0 & 0 & b & 0 \\ 0 & 0 & 0 & c \\ d & 0 & 0 & 0 \end{bmatrix}$, где a, b, c, d — положительные числа.

21. Вычислить индекс импримитивности неразложимых матриц из упражнения 20. Для примитивных матриц найти их индекс примитивности.

22. Выяснить, существует ли $\lim\limits_{m\to\infty} A^m$ для матрицы A:
a) $\begin{bmatrix} 0,3 & 0 & 0,2 & 0,5 \\ 0,5 & 0,2 & 0 & 0,3 \\ 0,4 & 0,3 & 0,3 & 0 \\ 0,9 & 0 & 0 & 0,1 \end{bmatrix}$, b) $\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0,5 & 0 & 0 & 0,5 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$.

23. Показать, что для любой унитарной матрицы $U \in \mathbb{C}^{n\times n}$ матрица $C \in \mathbb{C}^{n\times n}$: $c_{ij} = |u_{ij}|^2$ является двоякостохастической.

24. Доказать, что для двоякостохастической матрицы A верно $e^T A = e^T$, где $e^T = [1, 1, \ldots, 1]$.

25. Найти $L = \lim\limits_{m\to\infty} \left[\rho\,[A]^{-1} A\right]^m$ для матриц a) $\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$, b) $\begin{bmatrix} 1 & 2 \\ 3 & 1 \end{bmatrix}$.

26. Найти перронов корень и перронов вектор для положительной матрицы из задачи 25.

27. Доказать свойства неразложимых матриц:

1^0. Если $A \in \mathbb{R}^{n\times n}$, $A \geq 0$, — неразложимая матрица и $x \in \mathbb{R}^n$, $x \geq 0$, $x \neq 0$, то $Ax \neq 0$.

2^0. Пусть $A \in \mathbb{R}^{n\times n}$, $A \geq 0$, — неразложимая матрица и $x \in \mathbb{R}^n$, $x \geq 0$, — вектор с k нулями $(1 \leq k < n)$. Тогда число нулевых координат в векторе $z = [E_n + A]\,x$ меньше k.

28. Показать, что для примитивной стохастической матрицы $A \in \mathbb{R}^{n\times n}$ матрица $A^\infty = \lim\limits_{m\to\infty} A^m$ имеет вид
$$\begin{bmatrix} y_1 & y_2 & \dots & y_n \\ y_1 & y_2 & \dots & y_n \\ \dots & \dots & \dots & \dots \\ y_1 & y_2 & \dots & y_n \end{bmatrix},$$
где $\sum_{i=1}^n y_i = 1,\ A^T = y,\ y^T = (y_1, y_2, \dots, y_n)$.

29. Найти $A^\infty = \lim\limits_{m\to\infty} A^m$ для матрицы A из примера 9.11.

30. Доказать, что если матрица A примитивна и $A^p > 0$, то $A^m > 0$, для всех $m > p$.

31. Доказать, что степень примитивной матрицы всегда неразложима и притом примитивна.

32. Доказать, что неотрицательная матрица $\left[\begin{smallmatrix} a_{11} & a_{12} \\ a_{21} & a_{22}\end{smallmatrix}\right]$ примитивна тогда и только тогда, когда одновременно выполняются условия: 1) $a_{12}, a_{21} \neq 0$, 2) $a_{11} \neq 0$ или $a_{22} \neq 0$ и импримитивна тогда и только тогда, когда одновременно выполняются условия: 3) $a_{12}, a_{21} \neq 0$, 4) $a_{11} = 0, a_{22} = 0$.

33. Доказать, что любая неотрицательная 3×3 матрица либо примитивна, либо импримитивна с индексом импримитивности $k = 3$.

34*. Матрица $A \geq 0$ называется *продуктивной*, если $\forall y \geq 0\ \exists x \geq 0 :$ $x = Ax + y$. Доказать, что матрица $A \geq 0$ продуктивна тогда и только тогда, когда $\rho(A) < 1$.

35*. Доказать, что матрица $A \geq 0$ продуктивна тогда и только тогда, когда $B = (E - A)^{-1} \geq 0$.

36*. Доказать, что если матрица $A \geq 0$ продуктивна, то она является сходящейся.

37. Доказать, что для любой матрицы $A \geq 0$, в каждой строке которого есть хотя бы один ненулевой элемент, матрица DA, где $D = diag\left\{\left[\sum_j a_{ij}^{-1},\ i = \overline{1,n}\right]\right\}$, является стохастической.

Вопросы для повторения

1. Как понимается соотношение $A \leq B$ для $A, B \in \mathbb{R}^{n\times n}$?
2. Перечислите свойства неотрицательных матриц.
3. Как для произвольной матрицы $A \in \mathbb{R}^{n\times n}$ соотносятся спектральные радиусы $\rho(A)$ и $\rho(|A|)$?
4. Пусть $A, B \in \mathbb{R}^{n\times n}$ — неотрицательные матрицы, для которых справедливо $A \leq B$. Как соотносятся спектральные радиусы этих матриц?
5. Когда для неотрицательной матрицы A ее спектральный радиус совпадает с максимальной строчной нормой? Минимальной столбцовой нормой?
6. Какие оценки спектрального радиуса можно получить с помощью строчных и столбцовых сумм?
7. Что такое перронов вектор? Левый перронов вектор?
8. Сформулируйте теорему Фань Цзы. Как она связана с теоремами Гершгорина?
9. Как можно вычислить индекс импримитивности?
10. Сформулируйте критерии примитивности матриц.
11. Назовите основные свойства примитивных матриц.
12. Какой должна быть матрица $A \in \mathbb{C}^{1\times 1}$, чтобы быть примитивной?
13. Какая матрица называется стохастической? Двоякостохастической? Чему равны их спектральные радиусы?
14. Когда для неразложимой стохастической матрицы A существует $A^\infty = \lim\limits_{m\to\infty} A^m$? Как найти этот предел?
15. Может ли степень неразложимой матрицы быть разложимой?
16. Может ли степень разложимой матрицы быть неразложимой?

17. Может ли степень примитивной матрицы быть импримитивной?

Необходимо усвоить.

Определения и свойства матриц

- двоякостохастическая
- импримитивная
- неотрицательная
- неположительная
- положительная
- примитивная
- стохастическая
- устойчивая
- индекс
  – – импримитивности
  – – примитивности
- перронов вектор
- левый перронов вектор
- перронов корень
- число Фробениуса
- вектор Фробениуса

Формулировку и трактовку теорем

- теорема Перрона
- теорема Перрона— Фробениуса
- теорема Фань Цзы
- графовый критерий примитивности

- критерий стохастичности матрицы
- теорема Биркгофа

Основные умения и навыки

- определять принадлежность заданной матрицы определенному классу
- вычислять индекс примитивности и индекс импримитивности заданной матрицы
- находить перронов вектор (левый перронов вектор) и перронов корень заданной матрицы
- находить число Фробениуса и вектор Фробениуса заданной матрицы

Глава 10*

Вполне неотрицательные и оцилляционные матрицы

Рекомендуемая литература: [9, стр. 371]

Рассмотрим матрицы, у которых не только элементы, но и все миноры любых порядков неотрицательны. Такие матрицы имеют важные применения в теории малых колебаний линейных упругих систем.

Прямоугольная матрица $A \in \mathbb{R}^{m\times n}$ называется *вполне неотрицательной (вполне положительной)*, если все миноры любых порядков этой матрицы неотрицательны (соответственно положительны):

$$A\begin{bmatrix} i_1 & i_2 & \dots & i_p \\ k_1 & k_2 & \dots & k_p \end{bmatrix} \geq 0 \text{ (соответственно } > 0)$$

$$\begin{bmatrix} 1 \leq i_1 < i_2 < \dots < i_p \leq n \\ 1 \leq k_1 < k_2 < \dots < k_p \leq n \\ p = 1, 2, \dots, \min(m, n) \end{bmatrix}.$$

Пример 10.1 (Вполне неотрицательная матрица). Рассмотрим неотрицательную матрицу

$$A=\begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}.$$

Поскольку она неотрицательная, то, во-первых, все ее миноры первого порядка — неотрицательны:

$$A\left[\begin{smallmatrix}1\\1\end{smallmatrix}\right] = a_{11} = 1 > 0,\ A\left[\begin{smallmatrix}1\\2\end{smallmatrix}\right] = a_{12} = 2 > 0,$$
$$A\left[\begin{smallmatrix}2\\1\end{smallmatrix}\right] = a_{21} = 0,\ A\left[\begin{smallmatrix}2\\2\end{smallmatrix}\right] = a_{22} = 3 > 0.$$

Минор второго порядка этой матрицы

$$A\left[\begin{smallmatrix}1&2\\1&2\end{smallmatrix}\right] = \det A = 3 > 0,$$

поэтому матрица A является вполне неотрицательной.

Неотрицательная матрица

$$B=\begin{bmatrix} 1 & 3 \\ 2 & 0 \end{bmatrix}$$

имеет отрицательный минор второго порядка

$$B\left[\begin{smallmatrix}1&2\\1&2\end{smallmatrix}\right] = -6 < 0,$$

поэтому матрица B не является вполне неотрицательной.

Неотрицательная матрица

$$C=\begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 2 & 3 \end{bmatrix}$$

имеет миноры второго порядка

$$C\left[\begin{smallmatrix}1&2\\1&2\end{smallmatrix}\right] = 1 > 0,\ C\left[\begin{smallmatrix}1&2\\1&3\end{smallmatrix}\right] = 1 > 0,\ C\left[\begin{smallmatrix}1&2\\2&3\end{smallmatrix}\right] = 1 > 0,$$
$$C\left[\begin{smallmatrix}1&3\\1&2\end{smallmatrix}\right] = 2 > 0,\ C\left[\begin{smallmatrix}1&3\\1&3\end{smallmatrix}\right] = 3 > 0,\ C\left[\begin{smallmatrix}1&3\\2&3\end{smallmatrix}\right] = 3 > 0,$$
$$C\left[\begin{smallmatrix}2&3\\1&2\end{smallmatrix}\right] = 2 > 0,\ C\left[\begin{smallmatrix}2&3\\1&3\end{smallmatrix}\right] = 3 > 0,\ C\left[\begin{smallmatrix}2&3\\2&3\end{smallmatrix}\right] = 4 > 0,$$

и минор третьего порядка

$$C\begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ 1 & 2 & 3 \end{bmatrix} = 1 > 0,$$

поэтому матрица C является вполне неотрицательной.

Будем рассматривать квадратные вполне неотрицательные и вполне положительные матрицы.

Матрица $A \in \mathbb{R}^{n\times n}$ называется *осцилляционной,* если A — вполне неотрицательная матрица и существует такое целое число $q > 0$, что A^q — вполне положительная матрица.

Пример 10.2 (Осцилляционная матрица). Вполне неотрицательная матрица $A = \left[\begin{smallmatrix} 1 & 2 \\ 0 & 3 \end{smallmatrix}\right]$ из предыдущего примера является разложимой матрицей, поэтому для любого $q > 0$ матрица A^q не будет положительной, а значит, и вполне положительной матрицей. Следовательно, матрица A не является осцилляционной матрицей.

Для вполне неотрицательной матрицы C из предыдущего примера имеем

$$C^2 = \begin{bmatrix} 2 & 3 & 3 \\ 3 & 7 & 9 \\ 2 & 10 & 15 \end{bmatrix} > 0,$$

значит, миноры первого порядка матрицы C^2 положительны. Легко подсчитать, что миноры второго порядка и единственный минор третьего порядка матрицы C^2 также положительны, значит матрица C является осцилляционной матрицей.

Теорема 10.1. *[9, стр. 376] Для того чтобы вполне неотрицательная матрица A была осцилляционной, необходимо и достаточно, чтобы выполнялись условия:*

1) A — невырожденная матрица ($\det A \neq 0$);

2) все элементы матрицы A, расположенные на главной диагонали, на первой наддиагонали и на первой поддиагонали, отличны от нуля ($a_{ik} > 0$ при $|i-k| \le 1$).

Упражнение 10.1. Применить теорему 10.1 для анализа матриц из примеров 10.1, 10.2.

Для того чтобы сформулировать свойства собственных значений и собственных векторов осцилляционной матрицы, введем некоторые обозначения и понятия.

Рассмотрим вектор

$$u = (u_1, u_2, \ldots, u_n).$$

Будем подсчитывать число перемен знака в ряду кординат $u_1, u_2, \ldots, u_n$ вектора u, приписывая нулевым координатам (если таковые имеются) произвольные знаки. В зависимости от того, какие знаки мы припишем нулевым координатам, число перемен знака будет колебаться в известных пределах. Получающиеся при этом *максимальное* и *минимальное* числа перемен знака будем обозначать, соответственно, через S_u^+ и S_u^-. В том случае, когда $S_u^- = S_u^+$, будем говорить о *точном* числе перемен знака и обозначать его через S_u. Очевидно, что $S_u^+ = S_u^-$ тогда и только тогда, когда 1) крайние координаты u_1 и u_n вектора u_0 отличны от нуля, и 2) равенство $u_i = 0$ $(1 < i < n)$ всегда сопровождается неравенством $u_{i-1}u_{i+1} < 0$.

Пример 10.3 (Число перемен знака). Подсчитаем число перемен знака для вектора $u = (1, -2, 0, 3, 4, -2, -1, 0)$. Вектор знаков может иметь один из четырех видов: $(+,-,-,+,+,-,-,-)$, $(+,-,-,+,+,-,-,+)$ — число перемен знаков равно 3, $(+,-,+,+,+,-,-,-)$, $(+,-,+,+,+,-,-,+)$ — число перемен знаков равно 4. Таким образом, $S_u^- = 3$, $S_u^+ = 4$.

Теорема 10.2. *[9, стр. 376] а) Осцилляционная матрица A всегда имеет n различных положительных собственных значений:*

$$\lambda_1 > \lambda_2 > \cdots > \lambda_n > 0. \tag{10.1}$$

b) У собственного вектора $u^1 = (u_{11}, u_{21}, \dots, u_{n1})$ матрицы A, отвечающего наибольшему собственному значению λ_1, все координаты отличны от нуля и одного знака; у собственного вектора $u^2 = (u_{12}, u_{22}, \dots, u_{n2})$ матрицы A, отвечающего второму по величине собственному значению λ_2, в ряду координат имеется точно одна перемена знака, и вообще в ряду координат собственного вектора $u^k = (u_{1k}, u_{2k}, \dots, u_{nk})$ матрицы A, отвечающего собственному значению λ_k, имеется точно $k-1$ перемен знака $(k = 1, 2, \dots, n)$.

Пример 10.4 (Собственные значения и собственные векторы осцилляционной матрицы). Матрица

$$D = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix}$$

является вполне неотрицательной и, согласно теореме 10.1 является осцилляционной. Ее собственные значения равны $\lambda_1 = \frac{1}{2}(5 + \sqrt{17})$, $\lambda_2 = 2$, $\lambda_3 = \frac{1}{2}(5 - \sqrt{17})$, и все различны и положительны. Собственные векторы имеют вид $u^1 = (1, \frac{1+\sqrt{17}}{2}, 1)$, $u^2 = (-1, 0, 3)$, $u^3 = (1, \frac{1-\sqrt{17}}{2}, 1)$.

Задачи и упражнения

1. Трехдиагональная вещественная матрица

$$A = \begin{bmatrix} a_1 & b_1 & 0 & \dots & 0 & 0 \\ c_1 & a_2 & b_2 & \dots & 0 & 0 \\ 0 & c_2 & a_3 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & c_{n-1} & a_n \end{bmatrix} > 0,$$

порядка n называется *якобиевой,* если $a_i c_{i-1} > 0$ для $i = 2, \dots, n$. Доказать, что якобиева матрица является осцилляционной в том и только том случае, когда 1) все числа b, c положительны и 2) последовательные главные миноры положительны.

2. Доказать, что произведение вполне неотрицательных (вполне положительных) матриц есть вполне неотрицательная (вполне положительная) матрица.
3. Привести примеры неотрицательных матриц 3×3, которые

   a) являются вполне неотрицательными;

   b) не являются вполне неотрицательными;

   c) являются вполне положительными;

   d) не являются вполне положительными;

   e) являются вполне неотрицательными, но не являются осцилляционными;

   f) являются осцилляционными.

Вопросы для повторения

1. Какая матрица называется вполне неотрицательной? Вполне положительной?
2. Какая матрица называется осцилляционной матрицей?
3. Сформулируйте критерий осцилляционности вполне неотрицательной матрицы.

Необходимо усвоить.

Определения и свойства матриц

- вполне неотрицательная
- вполне положительная
- осцилляционная

Основные умения и навыки

- определять принадлежность заданной матрицы классу вполне положительных, вполне неотрицательных, осцилляционных матриц

ОТВЕТЫ И УКАЗАНИЯ

Ответы и указания к главе 1

1. $r(\lambda) = f(\mu) + \dot f(\mu)(\lambda-\mu) + \frac{\ddot f(\mu)}{2!}(\lambda-\mu)^2 + \ldots + \frac{f^{(n-1)}(\mu)}{(n-1)!}(\lambda-\mu)^{n-1}$,

   $f(A) = \begin{bmatrix} f(\mu) & \dot f(\mu) & \frac{\ddot f(\mu)}{2!} & \ldots & \frac{f^{(n-1)}(\mu)}{(n-1)!} \\ 0 & f(\mu) & \dot f(\mu) & \ldots & \frac{f^{(n-2)}(\mu)}{(n-2)!} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & f(\mu) \end{bmatrix}$. $f(A)$ имеет смысл для любой функции $f(\lambda)$, для которой существуют значения $f(\mu), \dot f(\mu), \ldots, f^{(n-1)}(\mu)$.

2. Применить результаты упражнения 1.4.

3. a) $\begin{bmatrix} 3\cdot 2^{100}-2\cdot 3^{100} & 2(3^{100}-2^{100}) \\ -3(3^{100}-2^{100}) & 3^{101}-2^{101} \end{bmatrix}$, b) $2^{50}\begin{bmatrix} -24 & 25 \\ -25 & 26 \end{bmatrix}$,

   c)$\pm\frac{1}{4}\begin{bmatrix} 7 & 1 \\ -1 & 9 \end{bmatrix}$, d)$\pm\frac{1}{5}\begin{bmatrix} 12 & 2 \\ 3 & 13 \end{bmatrix}$, $\pm\begin{bmatrix} 0 & 2 \\ 3 & 1 \end{bmatrix}$, всего четыре матрицы,

   e) $\begin{bmatrix} 2e^2 & -e^2 \\ e^2 & 0 \end{bmatrix}$, f) $\begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}$, g)$\begin{bmatrix} 3e-1 & e & -3e+1 \\ 3e & e+3 & -3e-3 \\ 3e-1 & e+1 & -3e \end{bmatrix}$,

   h) $\begin{bmatrix} 3 & -15 & 6 \\ 1 & -5 & 2 \\ 1 & -5 & 2 \end{bmatrix}$, если брать вещественное значение логарифма;

   i) $\frac{1}{4}\begin{bmatrix} 9 & -1 & 2 \\ 1 & 7 & 2 \\ 0 & 0 & 8 \end{bmatrix}$, j) $\begin{bmatrix} e^{-1} & e^2-e^{-1} & e^{-1}-e^2 \\ e-e^{-1} & e^{-1} & e-e^{-1} \\ e-e^{-1} & e^{-1}-e^2 & e^2-e^{-1}+e \end{bmatrix}$,

   l) $\begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix}$, m) $\begin{bmatrix} \cos\alpha & i\sin\alpha \\ i\sin\alpha & \cos\alpha \end{bmatrix}$, n) $\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$,

   o) $e^A = \begin{bmatrix} 6-7e & 3-4e & 2-3e \\ -6+10e & -3+6e & -2+4e \\ -6+6e & -3+3e & -2+3e \end{bmatrix}$, $\sqrt{A} = \begin{bmatrix} -6 & -3{,}5 & -2{,}5 \\ 8 & 5 & 3 \\ 6 & 3 & 3 \end{bmatrix}$,

   p) $\cos A = \cos(1)E$.

6. $\sin A = \begin{bmatrix} \sin 1 & \sin 1 & \sin 1 & \sin 1 \\ 0 & -\sin 1 & -2\sin 1 & -3\sin 1 \\ 0 & 0 & \sin 1 & 3\sin 1 \\ 0 & 0 & 0 & -\sin 1 \end{bmatrix}$.

7. $J_n^m(\mu) = \begin{bmatrix} \mu^m & \binom{m}{1}\mu^{m-1} & \ldots & \binom{m}{n-1}\mu^{m-n+1} \\ & \mu^m & \ldots & \binom{m}{n-2}\mu^{m-n+2} \\ & & \ldots & \ldots \\ & & & \mu^m \end{bmatrix}$.

8. 1b), 2b), 3a), 3b), 4a), 5a) — да, 1a), 2a), 4b), 5b) — нет.

Ответы и указания к главе 2

1. Рассмотреть AB как скелетное разложение некоторой матрицы, воспользоваться для нее (2.12), использовать свойства 7^0, 8^0 из теоремы 2.4.

5. (a) $B = \begin{bmatrix} 2i & i-1 \\ 0 & -2 \\ 3+i & 0 \end{bmatrix}$, $C = \begin{bmatrix} 1 & 0 & i & -i \\ 0 & 1 & 0 & 2+i \end{bmatrix}$;

   (b) $B = \begin{bmatrix} 1 & 0 \\ -1 & 2 \\ -1 & 0 \end{bmatrix}$, $C = \begin{bmatrix} 1 & 0 & -1i & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix}$;

   (c) $B = \begin{bmatrix} 1 & -2 \\ 0 & 2 \\ -1 & 3 \end{bmatrix}$, $C = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 2 & -1 \end{bmatrix}$,

   (d) $B = A, \quad C = E_2$.

6. Свойства a)— d) следуют непосредственно из определения псевдообратной матрицы. Свойства e), f) вытекают из (2.1), 5^0 и эрмитовости матриц AA^+ и A^+A (см. 3^0, 4^0, теорема 2.2). Для получения g) достаточно воспользоваться определением псевдообращения Мура—Пенроуза и равенствами e), f). Свойство h) следует из свойств f) и g); свойство i) вытекает из g) и e). Для доказательства j) заметим, что в силу i) выполнено равенство $B^+A^+ = (B^*B)^+ B^*A^* (AA^*)^+$. Чтобы доказать l), воспользуемся h) и i) и запишем следующие соотношения: $A^+B = 0 \Leftrightarrow (A^*A)^+ A^*B = 0 \Leftrightarrow A^*A(A^*A)^+ A^*B = 0 \Leftrightarrow$ $\Leftrightarrow A^*B = 0$. n) Рассмотрите $A = -E$.

16. a) $A^+ = \begin{bmatrix} \frac{2}{9} & \frac{1}{9} & \frac{-2}{9} \end{bmatrix}^T$, b) $B^+ = \begin{bmatrix} \frac{1}{13} & 0 & \frac{2}{13} & \frac{1}{13} & \frac{1}{13} \\ \frac{19}{312} & \frac{1}{48} & \frac{21}{208} & \frac{17}{208} & \frac{25}{624} \\ \frac{5}{156} & \frac{-1}{24} & \frac{11}{104} & \frac{-1}{104} & \frac{23}{312} \\ \frac{3}{104} & \frac{1}{16} & \frac{-1}{208} & \frac{19}{208} & \frac{-7}{208} \end{bmatrix}$,

    c) $C^+ = \begin{bmatrix} \frac{-21}{134} & \frac{-1}{134} & \frac{2}{67} & \frac{2}{67} & \frac{7}{67} \\ \frac{-31}{134} & \frac{-27}{134} & \frac{-13}{67} & \frac{-13}{67} & \frac{-12}{67} \end{bmatrix}$,

    d) $D^+ = \begin{bmatrix} \frac{-2}{125} & 0 & \frac{1}{125} & \frac{2}{125} & \frac{6}{125} & 0 \\ \frac{1}{125} & 0 & \frac{-1}{250} & \frac{-3}{250} & \frac{-3}{125} & 0 \end{bmatrix}$.

18. a) $x^0 = \begin{bmatrix} -2 & 2 & -4 \end{bmatrix}^T$, $r = \begin{bmatrix} 0 & 18 & -18 \end{bmatrix}^T$, $|r| = 18\sqrt{2}$,

    b) $x^0 = \begin{bmatrix} -3 & 0 & -3 \end{bmatrix}^T$, $r = \begin{bmatrix} 2 & 2 & 0 \end{bmatrix}^T$, $|r| = 2\sqrt{2}$,

c) $x^0 = \begin{bmatrix} -2 & 0 & -2 \end{bmatrix}^T$, $r = \begin{bmatrix} 0 & -8 & -8 \end{bmatrix}^T$, $|r| = 8\sqrt{2}$.

Ответы и указания к главе 3

1. 1) $\left[\begin{array}{ccc|cc|cc|c} 0&0&0&a&b&0&0&c\\ 0&0&0&0&a&0&0&0\\ 0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0\\ \hline d&e&f&0&0&g&h&0\\ 0&d&e&0&0&0&g&0\\ 0&0&d&0&0&0&0&0\\ \hline 0&0&0&k&r&0&0&m\\ 0&0&0&0&k&0&0&0\\ \hline 0&0&n&0&0&0&p&0 \end{array}\right]$; 2) $\left[\begin{array}{cc|cc|cc|c|c} 0&0&0&0&a&b&c&0\\ 0&0&0&0&0&a&0&0\\ 0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0\\ \hline 0&0&0&0&d&e&f&0\\ 0&0&0&0&0&d&0&0\\ 0&0&0&0&0&0&0&0\\ \hline 0&0&g&h&0&0&0&i\\ 0&0&0&g&0&0&0&0\\ \hline 0&0&0&j&0&0&0&k \end{array}\right]$;

3) $\left[\begin{array}{c|ccc|ccc|cc|cc|c} 0&0&0&0&0&0&0&a&b&c&d&0\\ 0&0&0&0&0&0&0&0&a&0&c&0\\ 0&0&0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0&0&0\\ \hline 0&e&f&g&h&i&j&0&0&0&0&k\\ 0&0&e&f&0&h&i&0&0&0&0&0\\ 0&0&0&e&0&0&h&0&0&0&0&0\\ \hline 0&0&y&m&0&n&p&0&0&0&0&q\\ 0&0&0&y&0&0&n&0&0&0&0&0\\ \hline 0&0&0&0&0&0&0&r&s&t&u&0\\ 0&0&0&0&0&0&0&0&r&0&t&0\\ \hline 0&0&0&v&0&0&w&0&0&0&0&x \end{array}\right]$; 4) $\left[\begin{array}{cccc|ccc|cc} a&b&c&d&0&0&0&e&f\\ 0&a&b&c&0&0&0&0&e\\ 0&0&a&b&0&0&0&0&0\\ 0&0&0&a&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0\\ \hline 0&0&0&0&g&h&i&0&0\\ 0&0&0&0&0&g&h&0&0\\ 0&0&0&0&0&0&g&0&0\\ 0&0&0&0&0&0&0&0&0\\ \hline 0&j&k&r&0&0&0&m&n\\ 0&0&j&k&0&0&0&0&m\\ 0&0&0&j&0&0&0&0&0\\ \hline 0&0&0&0&0&0&p&0&0 \end{array}\right]$;

5) $\left[\begin{array}{cc|ccc|ccc|cc|c} 0&0&a&b&c&d&e&f&0&0&0\\ 0&0&0&a&b&0&d&e&0&0&0\\ 0&0&0&0&a&0&0&d&0&0&0\\ \hline 0&0&0&0&0&0&0&0&g&h&i\\ 0&0&0&0&0&0&0&0&0&g&0\\ \hline 0&0&0&0&j&0&0&k&0&0&0\\ \hline 0&0&0&0&0&0&0&0&0&l&m \end{array}\right]$.

2. 1) 13. 2) 23. 3) 13. 4) 15. 5) 11.

3. 1) $X = \begin{bmatrix} -a+2(c+e) & a-c-e & a+b+d-2(+e)\\ -2a+2e & 2a-e & 2a+d-2e\\ 2c & -c & b-2c \end{bmatrix}$

2) $X = \begin{bmatrix} -2a+3b+4c & a-b-c & a-b-2c\\ -4a+4b & 2a-b & 2[a-b]\\ -2a+2b+8c & a-b-2c & a-4c \end{bmatrix}$

3) $X = \begin{bmatrix} a+3b+2c & -b-c & 2b+c\\ 5b+2c & a-2b-c & 3b+c\\ -b-2c & c & a-b-c \end{bmatrix}$

4) $X = \begin{bmatrix} 2a-b+c & -6a+6b-5c & -5a+5b-4c\\ 2a-2b-3c & -6a+7d+15c & -5a+5b+12c\\ -2a+2b+4c & 6a-6b-20c & 5a-4b-16c \end{bmatrix}$

5) $X = \begin{bmatrix} b+4c & 0 & 8c\\ \frac{3}{4}(-a+b-4c) & a & 6c\\ -2c & 0 & b-4c \end{bmatrix}$

6) $X = \begin{bmatrix} \frac{1}{2}[4a+6b-2c-3d+e] & -2a-5b+c+\frac{5d}{2}-\frac{e}{2} & 2a+6b-c-3d+e\\ \frac{1}{2}[-2c-3d+3e] & c+\frac{5d}{2}-\frac{3e}{2} & -c-3d+3e\\ -a-\frac{3b}{2}+e & a+\frac{5b}{2}-e & -a-3b+2e \end{bmatrix}$

4. 1) 83. 2) 65. 3) 66. 4) 152. 5) 129.

6. 1) Фундаментальная система решений:

$X_1 = \begin{bmatrix} 0&0&0\\ 1&-2&-1\\ 2&-4&-2 \end{bmatrix}$, $X_2 = \begin{bmatrix} -14&26&12\\ -7&13&6\\ 15&-28&-13 \end{bmatrix}$, $X_3 = \begin{bmatrix} 2&-1&-2\\ 1&-2&-1\\ -2&4&2 \end{bmatrix}$,

$X = aX_1 + bX_2 + cX_3, \quad a,b,c \in \mathbb{R}.$

2) Фундаментальная система решений:

$X_1 = \begin{bmatrix} 6&2\\ 3&1\\ -6&-2 \end{bmatrix}$, $X_2 = \begin{bmatrix} 3&1\\ 3&0\\ -6&-2 \end{bmatrix}$, $X = aX_1 + bX_2, \quad a,b \in \mathbb{R}.$

3) Фундаментальная система решений:

$X_1 = \begin{bmatrix} 18&-9&-9\\ 0&0&0\\ -6&3&3 \end{bmatrix}$, $X_2 = \begin{bmatrix} 30&-15&-15\\ -6&3&3\\ 0&0&0 \end{bmatrix}$, $X_3 = \begin{bmatrix} -1&-1&-1\\ -1&-1&-1\\ 3&3&3 \end{bmatrix}$,

$X = aX_1 + bX_2 + cX_3, \quad a,b,c \in \mathbb{R}.$

4) Фундаментальная система решений:

$X_1 = \begin{bmatrix} 0&0&0\\ -3&-15&-6\\ 3&15&6 \end{bmatrix}$, $X_2 = \begin{bmatrix} 0&0&0\\ 1&8&3\\ -1&-8&-3 \end{bmatrix}$,

$X_3 = \begin{bmatrix} -3&-15&-6\\ 6&30&12\\ -3&-15&-6 \end{bmatrix}$, $X_4 = \begin{bmatrix} 1&8&3\\ -2&-16&-6\\ 1&8&3 \end{bmatrix}$,

$X = aX_1 + bX_2 + cX_3 + dX_4, \quad a,b,c,d \in \mathbb{R}.$

5) Фундаментальная система решений:

$X_1 = \begin{bmatrix} -2&2&-1\\ -2&2&-1\\ -2&2&-1 \end{bmatrix}$, $X_2 = \begin{bmatrix} 2&-1&0\\ 0&0&0\\ 2&-1&0 \end{bmatrix}$, $X_3 = \begin{bmatrix} 1&-1&1\\ 1&-1&1\\ 0&0&0 \end{bmatrix}$,

$X = aX_1 + bX_2 + cX_3, \quad a,b,c \in \mathbb{R}.$

6) Фундаментальная система решений:

$X_1 = \begin{bmatrix} 4&2&0\\ 6&3&0 \end{bmatrix}$, $X_2 = \begin{bmatrix} 2&2&2\\ 3&3&3 \end{bmatrix}$, $X = aX_1 + bX_2, \quad a,b \in \mathbb{R}.$

7) $X = \begin{bmatrix} 0&0\\ 0&0\\ 0&0 \end{bmatrix}$. 8) $X = \begin{bmatrix} 0&0&0\\ 0&0&0\\ 0&0&0 \end{bmatrix}$.

7. Использовать блочную структуру матрицы B и следствие 3.1.

Ответы и указания к главе 4

1. 1) биортогональный базис: $f_1(x,y,z) = \frac{1}{3}(x-y-z)$ $f_2(x,y,z) = \frac{1}{12}(2x+4y+z)$, $f_3(x,y,z) = \frac{1}{12}(2x+4y+7z)$, $f = 11f_1 - 14f_2 + 10f_3$, $u = -\frac{2}{3}e_1 + \frac{5}{12}e_2 + \frac{35}{12}e_3$.

2. 1) $W^{\perp} = L(f_1, f_2)$, $f_1(x,y,z,t,u) = 3x + 5y + 2z + t$, $f_2(x,y,z,t,u) = 2x + 3y + 7z + u$, $\dim \mathrm{W}^{\perp} = 2$.

3. 1) $\mathcal{A}^*(g)(x,y,z) = x - y$, 2) $\mathcal{A}^*(g)(x,y) = x - 4y$.

4. 1) $E^{\perp}: e_1^{\perp}(x,y) = \frac{1}{5}(2x-y)$, $e_2^{\perp}(x,y) = \frac{1}{5}(3x+y)$;
$H^{\perp}: h_1^{\perp}(x,y,z) = (-x-y+2z)$, $h_2^{\perp}(x,y,z) = (-x+2z)$, $h_3^{\perp}(x,y,z) = z$; $A_{E^{\perp}}^{H^{\perp}} = \begin{bmatrix} 9 & 5 & 3 \\ -6 & -5 & -2 \end{bmatrix}$.

5. Аналогично доказательству теоремы 4.4 с изменениями, вызванными заменой линейных функций на полулинейные функции.

Ответы и указания к главе 5

1. Использовать определение ортонормированной системы векторов.

3. От противного, используя упражнение 2.

5. a) $Q = \begin{bmatrix} \frac{3}{\sqrt{13}} & \frac{4}{13} \\ -\frac{2}{\sqrt{13}} & \frac{6}{13} \end{bmatrix}$, $R = \begin{bmatrix} \sqrt{13} & \frac{3}{\sqrt{13}} \\ 0 & 1 \end{bmatrix}$;

b) $Q = \begin{bmatrix} -\frac{2}{\sqrt{6}} & \sqrt{\frac{6}{35}} \times \frac{4}{3} & \frac{1}{\sqrt{35}} \\ -\frac{1}{\sqrt{6}} & -\frac{11}{\sqrt{210}} & \frac{3}{\sqrt{35}} \\ \frac{1}{\sqrt{6}} & \sqrt{\frac{5}{42}} & \sqrt{\frac{5}{7}} \end{bmatrix}$, $R = \begin{bmatrix} \sqrt{6} & \frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{6}} \\ 0 & \sqrt{\frac{35}{6}} & 8\sqrt{\frac{2}{105}} + \sqrt{\frac{5}{42}} + \frac{11}{\sqrt{210}} \\ 0 & 0 & \frac{4}{\sqrt{35}} \end{bmatrix}$;

c) $Q = \begin{bmatrix} \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{30}} & \frac{2}{\sqrt{5}} \\ \sqrt{\frac{2}{3}} & -\sqrt{\frac{2}{15}} & -\frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{6}} & \sqrt{\frac{5}{6}} & 0 \end{bmatrix}$, $R = \begin{bmatrix} \sqrt{6} & 5\sqrt{\frac{2}{3}} + \sqrt{\frac{3}{2}} & -2\sqrt{\frac{2}{3}} + \sqrt{\frac{3}{2}} - \frac{1}{\sqrt{6}} \\ 0 & \sqrt{\frac{5}{6}} & 2\sqrt{\frac{2}{15}} - \sqrt{\frac{3}{10}} - \sqrt{\frac{5}{6}} \\ 0 & 0 & \frac{8}{\sqrt{5}} \end{bmatrix}$.

7. a) непосредственно использовать определение евклидовой длины вектора; в) из теоремы о спектральном разложении нормальных матриц; геометрически это означает, что угол между векторами Ax и Ay тот же самый, что и между векторами A^*x и A^*y для всех $x, y \in \mathbb{C}^n$.

9. Использовать равносильность $x^*x = 0 \Leftrightarrow x = 0$.

10. Установить, что $U\Lambda U^*$ и $V\Lambda V^*$ унитарно подобны, если U и V унитарные. Привести пример двух (не являющихся нормальными) матриц, которые подобны, но не являются унитарно подобными.

11. Пусть $U \in \mathbb{C}^{n\times n}$ – унитарная матрица с первым столбцом, совпадающим с собственным вектором для A (и, следовательно, для A^*). Рассмотреть вместе U^*AU и $U^*A^*U = (U^*AU)^*$ и продолжить доказательство.

12. Записать $A = U\Lambda U^*$, где $\Lambda = \operatorname{diag}[\lambda_1, ..., \lambda_n]$ и $U \in \mathbb{C}^{n\times n}$ – унитарная матрица. Пусть $\chi = U^*x = (\chi_i)$ и $\eta = U^*y = (\eta_i)$. Показать, что $\Lambda\chi = \lambda\chi$, и отсюда вывести, что $\chi_i = 0$ для любого индекса i, такого, что $\lambda_i \neq \lambda$; то же проделать с η. Установить ортогональность χ и η, отсюда вывести ортогональность x и y.

13. a) $U = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$, $\Lambda = \begin{bmatrix} 1-i & 0 \\ 0 & 1+i \end{bmatrix}$; c) $U = \begin{bmatrix} \frac{\sqrt{3}}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix}$, $\Lambda = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$.

16. Пусть $U \in \mathbb{C}^{n\times n}$ — произвольная унитарная матрица. Тогда матрица U^*AU унитарно подобна матрице A. Далее доказательство вытекает из цепочки равносильностей:

A — нормальная $\Leftrightarrow$ (по определению) $AA^* = A^*A \Leftrightarrow$ $\Leftrightarrow U^*AA^*U = U^*A^*AU \Leftrightarrow$ (так как $UU^* = E$) $\Leftrightarrow U^*AUU^*A^*U = U^*A^*UU^*AU \Leftrightarrow U^*AU(U^*AU)^* =$ $= (U^*AU)^*U^*AU \Leftrightarrow$(по определению) U^*AU — нормальная.

Ответы и указания к главе 6

2. $A = H + iT$, где $H = \begin{bmatrix} 0 & -1-\frac{3}{2}i \\ -1+\frac{3}{2}i & 1 \end{bmatrix}$, $\quad T = \begin{bmatrix} 1 & \frac{3}{2}+i \\ \frac{3}{2}-i & 0 \end{bmatrix}$.

3. Использовать теорему 5.6 и теорему 6.2 b).

4. Использовать свойство 7^0 эрмитовых матриц и теорему 6.2 b).

5. Аналогично доказательству положения d) теоремы 5.7 стр. 155.

6. Использовать упражнение 5 стр. 199. $A = P\Lambda P^T$, где $P = \begin{bmatrix} \frac{2}{3} & \frac{2}{3} & -\frac{1}{3} \\ \frac{2}{3} & -\frac{1}{3} & \frac{2}{3} \\ -\frac{1}{3} & \frac{2}{3} & \frac{2}{3} \end{bmatrix}$, $\Lambda = \begin{bmatrix} 3 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 9 \end{bmatrix}$.

7. Использовать последовательно определения унитарной и эрмитовой матрицы.

8. Использовать соотношение (6.5).

9. Использовать определение собственного вектора и собственного значения матриц.

11. Положить вектор x равным x_i, где x_i — собственный вектор матрицы A, соответствующий собственному значению λ_i и использовать определение собственного значения: $Ax_i = \lambda_i x_i$.

14. Аналогично примеру 6.5 показать, что в силу равенства смешанных частных производных симметричная матрица $\frac{1}{2}\left[A + A^T\right]$ приводит к тому же оператору, что и матрица A.
17. a), b) — да, c) — нет, d), e) — да.
18. Пользуясь разложением Такаги, свойствами комплексного сопряжения, определением унитарной матрицы U, а также тем, что матрица Σ вещественная, показать, что матрица $A\bar{A}$ унитарно подобна диагональной матрице $\Sigma^2 = \operatorname{diag}\left[\sigma_1^2, ..., \sigma_n^2\right]$, а значит, имеет те же собственные значения.
19. Если матрицы $B, C \in \mathbb{C}^{n\times n}$ невырожденные, то $\operatorname{rank} A = \operatorname{rank} BAC$.
20. Непосредственно установить, что $A\bar{A} = S\Delta S^{-1}$, где $\Delta = \operatorname{diag}\left\{|\lambda_1|^2, ..., |\lambda_n|^2\right\}$; использовать указание к упражнению 19.
21. Использовать равенство $x^T(Ay) = (Ax)^T y$.
24. , 25. Использовать теорему 5.6.
26. Использовать теорему Рэлея—Ритца и определение положительно (неотрицательно) определенной матрицы.
27. Если $\lambda_1, ..., \lambda_n$ – собственные значения матрицы A, то $\lambda_1^k, ..., \lambda_n^k$ – это собственные значения матрицы A^k. Далее использовать предыдущее упражнение.
28. Применить к рассматриваемой матрице определение положительно полуопределенной матрицы.
29. A, F, G — да, B, C, D — нет.
31. Доказывается непосредственно с использованием (2.7) стр. 39.
32. Использовать определение обратной матрицы.
33. Использовать разложения Такаги для симметричной матрицы A и представление $\Sigma^{\frac{1}{2}} = \operatorname{diag}\left[\sqrt{\sigma_1}, \sqrt{\sigma_2}, \dots, \sqrt{\sigma_n}\right]$, $S = U\Sigma^{\frac{1}{2}}$.
34. Проверить, что матрица A не является эрмитовой.

    Рассмотреть функцию

$$g(x) = \frac{x^TAx}{x^Tx} = \frac{x_1^2 + 2x_1x_2 + x_2^2}{x_1^2 + x_2^2},$$

определенную на векторах

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2, \; x \neq 0.$$

Для этой функции найти минимум и максимум. Убедится, что

$$\lambda_{\min} = \lambda_1 = 1 > \min_{x\neq 0} \frac{x^TAx}{x^Tx} = 0,$$

$$\lambda_{\max} = \lambda_2 = 1 < \max_{x\neq 0} \frac{x^TAx}{x^Tx} = 2,$$

т.е. соотношения (6.7), (6.8) теоремы Рэлея—Ритца не выполняются для матрицы A, не являющейся эрмитовой.

36. Непосредственно пользуясь определениями унитарных, эрмитовых и косоэрмитовых матриц показать, что если $U \in \mathbb{C}^{n\times n}$, $V \in \mathbb{C}^{n\times n}$ – унитарные, $A \in \mathbb{C}^{n\times n}$ – эрмитова, $B \in \mathbb{C}^{n\times n}$ косоэрмитова матрицы, то V^*UV – унитарная, V^*AV – эрмитова, V^*BV – косоэрмитова.
38. Для каждой матрицы M сингулярное разложение представлено в виде $M = U_M\Sigma_M V_M^*$, полярное разложение — в виде $M = W_M H_M$:

    $U_A = \frac{1}{15\sqrt{2}}\left[\begin{smallmatrix} 21 & 3 \\ 3 & -21 \end{smallmatrix}\right]$, $\Sigma_A = \left[\begin{smallmatrix} 15 & 0 \\ 0 & 5 \end{smallmatrix}\right]$, $V_A = \frac{1}{\sqrt{2}}\left[\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix}\right]$,

    $W_A = \frac{1}{5}\left[\begin{smallmatrix} 4 & 3 \\ -3 & 4 \end{smallmatrix}\right]$, $H_A = \left[\begin{smallmatrix} 10 & 5 \\ 5 & 10 \end{smallmatrix}\right]$;

    $U_B = \frac{1}{\sqrt{2}}\left[\begin{smallmatrix} 1+i & 0 \\ 0 & i-1 \end{smallmatrix}\right]$, $\Sigma_B = \left[\begin{smallmatrix} 2 & 0 \\ 0 & 2 \end{smallmatrix}\right]$, $V_B = \frac{1}{\sqrt{2}}\left[\begin{smallmatrix} 1 & i \\ i & 1 \end{smallmatrix}\right]$;

    $U_C = \frac{1}{2}\begin{bmatrix} \sqrt{3} & 1 \\ -1 & \sqrt{3} \end{bmatrix}$, $\Sigma_C = \left[\begin{smallmatrix} 3 & 0 \\ 0 & 1 \end{smallmatrix}\right]$, $V_C = \frac{1}{2}\begin{bmatrix} 1 & \sqrt{3} \\ -\sqrt{3} & 1 \end{bmatrix}$,

    $U_G = \frac{1}{\sqrt{2}}\left[\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix}\right]$, $\Sigma_G = \begin{bmatrix} \sqrt{45} & 0 \\ 0 & \sqrt{5} \end{bmatrix}$, $V_G = \frac{1}{\sqrt{10}}\left[\begin{smallmatrix} 3 & 1 \\ 1 & -3 \end{smallmatrix}\right]$,

    $W_G = \frac{1}{\sqrt{5}}\left[\begin{smallmatrix} 2 & -1 \\ 1 & 2 \end{smallmatrix}\right]$, $H_G = \frac{1}{\sqrt{5}}\left[\begin{smallmatrix} 14 & 3 \\ 3 & 6 \end{smallmatrix}\right]$.

Ответы и указания к параграфу 7.1

1. a) Использовать очевидное равенство $y = x + (y - x)$, $\forall x, y \in V$, свойства 1) — 3) векторной полунормы и определение функции модуль.

   b) Использовать $\|x - y\| = \|x \pm z \pm u - y\|$ $\forall x, y, u, z \in V$, и свойства 1) – 3) векторной полунормы.

2. Использовать *неравенство Минковского*
$$\left[\sum_{j=1}^{n}\left[x_j+y_j\right]^p\right]^{\frac{1}{p}} \le \left[\sum_{j=1}^{n} x_j^p\right]^{\frac{1}{p}} + \left[\sum_{j=1}^{n} y_j^p\right]^{\frac{1}{p}},$$
$x_j, y_j > 0,\ p \ge 1$.

3. a), b), c) — да, d)— нет.

5. Использовать свойства функции модуль. Для опровержения свойства 1a) векторной нормы достаточно рассмотреть функцию $f(x) = x - x_0 \in C\,[a,b]$.

10. Взять в качестве T в теореме 7.2 матрицу $T = \left[\begin{smallmatrix} 2 & -3 \\ 0 & 1 \end{smallmatrix}\right]$.

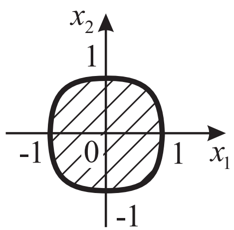
11. $\|x\|_\infty \le \|x\|_1 \le n\,\|x\|_\infty$, векторы – как в примере 7.5.

12. Воспользоваться определением шара при $x = 0,\ r = 1$.

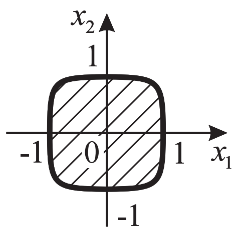
13. Рассмотреть (7.3) при $y = x + z,\ z \in V$.

14. $l_1 \subseteq l_2 \subseteq l_\infty$.

15. См. рисунок на стр.**??**



l_3



l_6

17. Сжимается в C раз.

18. $\|x\|_1 = 11, \|x\|_2 = \sqrt{35}, \|x\|_\infty = 4;\ \|y\|_1 = 9,\ \|y\|_2 = 5,$ $\|y\|_\infty = 4;\ \|z\|_1 = 5 + 2\sqrt{2}, \|z\|_2 = \sqrt{21},\ \|z\|_\infty = 4;\ \|u\|_1 = 9,$ $\|u\|_2 = 5, \|u\|_\infty = 4; \|v\|_1 = 3 + 2\sqrt{2},\ \|v\|_2 = 3, \|v\|_\infty = 2.$

20. Для доказательства неравенства треугольника преобразовать выражение $\|x+y\|^2$ и использовать *неравенство Коши—Шварца*: $|\langle x,y\rangle|^2 \le \langle x,x\rangle\,\langle y,y\rangle$.

21. Использовать тождество параллелограмма (7.4).

22. Использовать определение отрезка с концами в точках x_1 и $x_2:\ \alpha x_1 + (1-\alpha)x_2,\ \alpha \in [0,1]$.

Ответы и указания к параграфу 7.2

1. a) использовать равенство $B = A + (B - A), \quad \forall A, B \in \mathbb{C}^{n\times m}$ и аксиому 3); b) использовать аксиому 4).

2. Рассмотреть, например, векторную l_∞-норму и матрицу $A = \left[\begin{smallmatrix} 1 & -0{,}1 \\ -0{,}1 & 0 \end{smallmatrix}\right]$.

3. Использовать (7.14).

4. Проверить для $\|A\|_S$ выполнение аксиом 1) – 4) матричной нормы.

5. Для 1*, 2* использовать свойства функции max: $a_i \le b_i, \forall i \Rightarrow$ $\Rightarrow \max\limits_i a_i \le \max\limits_i b_i;\ \max\limits_i(a_i + b_i) \le \max\limits_i a_i + \max\limits_i b_i$; 4* Доказать и использовать, что если λ – собственное значение матрицы A, то $c\lambda$ – собственное значение матрицы cA; для 5* использовать очевидное неравенство $|a_k| \le \max\limits_k |a_k|$; для 6* использовать неравенство $\sum\limits_{k=1}^{n} a_{ik}b_{kj} \le \sum\limits_{k=1}^{n}\sum\limits_{s=1}^{n} a_{ik}b_{sj}$.

6. Использовать определение евклидовой и M-норм.

7. Рассмотреть матрицу $A = \left[\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right]$.

9. $\|A\|_1 = \frac{1}{\sqrt{3}} + \sqrt{\frac{8}{3}} \approx 2,21,\ \ \|A\|_\infty = \frac{4}{\sqrt{3}} \approx 2,31,\ \ \|A\|_E = \sqrt{6} \approx 2,45,$ $\|A\|_2 = 2,\quad \|A\|_M = 2\sqrt{3} \approx 3,46,\quad \|A\|_{l_1} = \frac{4+\sqrt{8}}{\sqrt{3}} \approx 3,94;$ $\|B\|_1 = 1 + \sqrt{5},\ \|B\|_\infty = 1 + \sqrt{5},\ \|B\|_E = \sqrt{7},\ \|B\|_2 = \sqrt{\frac{7+\sqrt{45}}{2}},$ $\|B\|_M = 2\sqrt{5},\ \|B\|_{l_1} = 2 + \sqrt{5};\ \|C\|_1 = 2 + \sqrt{5},\ \|C\|_\infty = 1 + \sqrt{5},$ $\|C\|_E = \sqrt{10},\ \|C\|_2 = \sqrt{5 + \sqrt{21}},\ \|C\|_M = 2\sqrt{5},\ \|C\|_{l_1} = 3 + \sqrt{5};$ $\|D\|_1 = 4,\ \|D\|_\infty = 3,\ \|D\|_E = \sqrt{10},\ \|D\|_2 = \sqrt{10},\ \|D\|_M = 4,$ $\|D\|_{l_1} = 6;\ \|F\|_1 = 1,\ \|F\|_\infty = 1,\ \|F\|_E = \sqrt{3},\ \|F\|_2 = 1,$ $\|F\|_M = 3,\ \|F\|_{l_1} = 3;\ \|G\|_1 = 10,\ \|G\|_\infty = 10,\ \|G\|_E = 9,$ $\|G\|_2 = 9,\ \|G\|_M = 12,\ \|G\|_{l_1} = 25;\ \|H\|_1 = 9;\ \|H\|_\infty = 12,$ $\|H\|_E = \sqrt{78},\ \|H\|_2 = 8.1737,\ \|H\|_M = 15,\ \|H\|_{l_1} = 24.$

11. a) во-первых, пользуясь определением скалярного произведения, неравенством Коши—Шварца и тем, что матричная спектральная норма подчинена евклидовой векторной норме, доказать, что для единичных векторов x и y справедливо неравенство $|y^*Ax| \le \|A\|_2$. Затем показать, что существуют такие векторы x и y единичной длины, на которых максимум достигается, взяв в качестве x такой вектор, что $\|Ax\|_2 = \|A\|_2$, т.е. на нем достигается максимум в определении подчиненной нормы, а в качестве $y = \frac{Ax}{\|Ax\|_2}$; b) использовать свойство a) и определение скалярного произведения; c) используя кольцевое свойство нормы и свойство b) спектральной нормы доказать справедливость неравенства $\|AA^*\|_2 \le \|A\|_2^2$. Затем, взяв такой вектор x, что $\|x\|_2 = 1$ и $\|Ax\|_2 = \|A\|_2$, положив $y = x$, получить неравенство $\|AA^*\|_2 \ge \|A\|_2^2$; d) использовать тот факт, что множество собственных значений блочно-диагональной матрицы есть совокупность собственных значений блоков; e) использовать свойства b) и c) спектральной нормы.

12. Аналогично примеру 7.11 подобрать такие матрицы, на которых достигаются неравенства.

13. $\sqrt{n}$.

16. $\rho(A^k) = 0,5^k$, $k = 1, 2, ...,$. $\lim\limits_{k\to\infty} \|A^k\|_1 = \lim\limits_{k\to\infty} \|A^k\|_\infty = \lim\limits_{k\to\infty} \|A^k\|_2 = 0$.

17. Выразить $x^{(k+1)}$ через $x^{(0)}$, $k = 0, 1, ...,$ и использовать (7.16).

18. Воспользоваться определением спектрального радиуса и оценкой (7.15).

19. Использовать определение (7.20) числа обусловленности.

20. a) использовать кольцевое свойство матричной нормы; b) нет.

21. Использовать эквивалентность матричных норм.

22. Использовать определение спектрального радиуса, неравенство (7.15) и $\lambda(A^{-1}) = \lambda^{-1}(A)$.

23. Рассмотреть матрицы вида $A = \varepsilon E_n$.

24. a) $\frac{\|x-\tilde{x}\|}{\|x\|} \le \frac{cond(A)}{1-\|A^{-1}\delta A\|}\frac{\|\delta A\|}{\|A\|}$; b) $\frac{\|x-\tilde{x}\|}{\|x\|} \le cond(A)\frac{\|\delta b\|}{\|b\|}$.

25. Использовать оценки (7.21) и (7.26).

26. $\mathrm{cond}_1 A = \frac{1}{2}(7+3\sqrt{5})$, $\mathrm{cond}_\infty A = (1+\frac{\sqrt{5}}{2})(2+\sqrt{5})$, $\mathrm{cond}_E A = 5$, $\mathrm{cond}_M A = \frac{45}{2}$, $\mathrm{cond}_{l_1} A = 7+3\sqrt{5}$, $\mathrm{cond}_2 A = \frac{1}{2}(5+\sqrt{21})$, $\mathrm{cond}_2 B = \frac{1}{2}(7+3\sqrt{5})$.

27. Рассмотреть евклидову норму единичной матрицы и использовать свойство (7.12) для подчиненной матричной нормы.

28. Использовать теорему 5.6, определение спектрального радиуса, соотношение между спектральным радиусом и нормами, а также определение M-нормы.

30. Напосредственно по определению спектральной нормы, сингулярных чисел и унитарной матрицы.

31. Использовать определения унитарно инвариантной нормы, числа обусловленности и свойства унитарных матриц.

33. Использовать упражнение 30. $cond_{\|\cdot\|_2}U$=1. Унитарная матрица идеально обусловлена по отношению к спектральной норме.

Ответы и указания к параграфу 8.1

1. В таблицах приведены центры z_i и радиусы R_i и C_i кругов строчных и столбцовых областей Гершгорина. Оценки спектрального радиуса получены из 8.3

a)
$\rho(A) \le 4$

i	z_i	R_i	C_i
1	0	2	3
2	$2i$	3	2

b)
$\rho(A) \le 5$

i	z_i	R_i	C_i
1	1	2	1
2	4	1	2

c)
$\rho(A) \le 4$

i	z_i	R_i	C_i
1	i	2	3
2	1	3	2

d)
$\rho(A) \le 3$

i	z_i	R_i	C_i
1	-1	2	1
2	$2i$	1	2

e)
$\rho(A) \le 31$

i	z_i	R_i	C_i
1	7	24	24
2	7	24	24
3	-5	16	16

f)
$\rho(A) \le 8.5$

i	z_i	R_i	C_i
1	0	2	1,5
2	6	3	2,5
3	5	1,5	1,5
4	-2	1	2

3. a) $\lambda_1 = i + \sqrt{5}$, $\lambda_2 = i - \sqrt{5}$, b) $\lambda_1 = 2$, $\lambda_2 = 3$,
c)$\lambda_1 = -2i$, $\lambda_2 = 3i$, d) $\lambda_1 = 1$, $\lambda_2 = 1$, e) $\lambda_1 = \lambda_2 = 7, \lambda_3 = -5$,
f) $\lambda_1 \approx 6,49 + 0,64i$, $\lambda_2 \approx 4,73 - 0,65i$, $\lambda_3 \approx -2,08 + 0,04i$, $\lambda_4 \approx -0,14 - 0,03i$.

4. Рассмотреть в оценке (7.15) нормы $\|\cdot\|_\infty$ и $\|\cdot\|_1$.

5. Применить теорему Гершгорина для $D^{-1}AD$ с учетом того факта, что для любой обратимой матрицы D матрица $D^{-1}AD$ имеет те же собственные значения, что и A.

6. (a) $\tilde{D} = \mathrm{diag}(p_1, p_2)$, где $\frac{p_1}{p_2} = \frac{a-d+\sqrt{(d-a)^2+4bc}}{2c}$,
(b) $r = \frac{a+d+\sqrt{(d-a)^2+4bc}}{2}$, (c), (d) $\rho(A) = r$.

7. (a) $\rho(A) = \frac{\sqrt{14}}{2} \approx 1,87$,
(в) $\min \left\| D^{-1}AD \right\|_\infty = \frac{3+\sqrt{7}}{2} \approx 2,82 > \rho(A)$.

8. (a) использовать $\rho(A) = \rho(D^{-1}AD)$ и (8.3).

10. Строчная область Гершгорина: $|z-1| \le 4$ и $|z-6| \le 1, \rho(A) \le 7$. Столбцовая область Гершгорина: $|z-1| \le 1$ и $|z-6| \le 4$, $\rho(A) \le 10$. Область Островского при $\alpha = \frac{1}{2}$: $|z-1| \le 2$ и $|z-6| \le 2$, $\rho(A) \le 8$.

11. a) $|a_{ii}| > R_i^\alpha C_i^{1-\alpha}$ для некоторого $\alpha \in [0, 1]$;
b) $|a_{ii}||a_{jj}| > R_i R_j$ для некоторых $1 \le i,\ j \le n,\ i \ne j$.

12. $\bigcup\limits_{i=1}^{n}\left\{z \in\mathbb{C}\colon |z - a_{ii}| \le\left(\frac{1}{p_i} \sum\limits_{\substack{j=1\\ j\ne i}}^{n} p_j|a_{ij}|\right)^{\alpha}\left(p_j \sum\limits_{\substack{j=1\\ j\ne i}}^{n} \frac{1}{p_i}|a_{ij}|\right)^{1-\alpha}\right\}$.

13. Ноль не принадлежит ни одному кругу Гершгорина. Можно также сказать, что A невырождена, т.к. она является матрицей со строгим диагональным преобладанием.

14. a) $|\lambda_i| \le 6,\ |\mathrm{Re}\lambda_\mathrm{i}| \le 5,\ |\mathrm{Im}\lambda_\mathrm{i}| \le 4,\ i = \overline{1,2}$;
b) $|\lambda_i| \le 8,\ |\mathrm{Re}\lambda_\mathrm{i}| \le 8,\ |\mathrm{Im}\lambda_\mathrm{i}| \le 3,\ i = \overline{1,2}$;
c) $|\lambda_i| \le 6,\ |\mathrm{Re}\lambda_\mathrm{i}| \le 1,\ |\mathrm{Im}\lambda_\mathrm{i}| \le 5,\ i = \overline{1,2}$;
d) $|\lambda_i| \le 4,\ |\mathrm{Re}\lambda_\mathrm{i}| \le 8,\ |\mathrm{Im}\lambda_\mathrm{i}| \le 2,\ i = \overline{1,2}$;
e) $|\lambda_i| \le 48,\ |\mathrm{Re}\lambda_\mathrm{i}| \le 48,\ |\mathrm{Im}\lambda_\mathrm{i}| \le 0,\ i = \overline{1,3}$;
f) $|\lambda_i| \le 24,\ |\mathrm{Re}\lambda_\mathrm{i}| \le 24,\ |\mathrm{Im}\lambda_\mathrm{i}| \le 2\sqrt{2},\ i = \overline{1,4}$.

15. Использовать матрицу $D = \mathrm{diag}\{1,\ 1,\ 1,\ 4\}$ и области Гершгорина для DAD^{-1}.

16. Собственные значения вещественных матриц образуют пары комплексно-сопряженных чисел, а изолированный круг Гершгорина содержит единственное собственное значение (см. пример 8.8).

17. По тем же причинам, что и 16.

18. С помощью теоремы Ольги Тауски; матрица A неразложима, является матрицей с диагональным преобладанием, причем для всех строк, кроме третьей, это свойство выполнено в сильной форме.

19. a) $\{-1,\ 1\}$, b) $[1-\sqrt{2}, 4+\sqrt{2}]$, c) $[(-1+\sqrt{2})i, (4+\sqrt{2})i]$, d) $[-(4+\sqrt{2})i, 7i]$, e) $[-7, 4+\sqrt{2}]$.

20. Узлы p_i и p_j обменяются номерами. Связность графа при этом не изменится.

21. Использовать очевидное равенство $\sum\limits_{i=1}^{n} R_i = \sum\limits_{k=1}^{n} C_k$.

22. От противного, используя обратимость матрицы со строгим диагональным преобладанием.

23. Показать, что в этом случае $\Gamma(A)$ не является сильно связным.

24. Использовать оценку (8.3).

27. Воспользоваться тем, что если $\det C \ne 0$, то A и AC одновременно являются невырожденными. Рассмотреть в качестве C матрицу перестановок и применить условия Адамара к AC.

28. Рассмотреть матрицу

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

30. Использовать область Гершгорина и вещественность собственных значений эрмитовой матрицы.

Ответы и указания к параграфу 8.2

1,2. Использовать теорему 8.15 и замечание к ней.

3. Рассмотреть область Гершгорина для матрицы $D + \delta D$.

4. Использовать унитарную диагонализуемость нормальных матриц (теорема 5.7) и упражнение 33 стр. 252.

5. $r = (0,0023;\ 0,00799;\ 0,05897)$, $\frac{\|r\|_2}{\|\tilde{x}\|_2} = 0,0601414$.

6. Использовать теорему Вейля, определение спектрального радиуса, утверждения c) теоремы 6.2, b) спектральной теоремы для нормальных матриц.

Ответы и указания к главе 9

1. (g)*. $\left|(AB)_{ij}\right| = \left|\sum\limits_{k=1}^{n} a_{ik}b_{kj}\right| \le \sum\limits_{k=1}^{n} |a_{ik}|\ |b_{kj}| = (|A|\ |B|)_{ij}$.

   (h)*. Из (g)* по индукции.

3. Наряду с главной подматрицей $\tilde{A}$ рассмотреть $n \times n$-матрицу $\widehat{A}$, которая получена из A заменой на нули всех элементов, не входящих в $\tilde{A}$, а элементы $\tilde{A}$ остаются неизменными на своих местах. Сравнить спектральные радиусы матриц $\tilde{A}$ и $\widehat{A}$, матрицы $\widehat{A}$ и A, затем воспользоваться следствием 9.1.

4. $\rho(A) = 2$, $A_1 = [1]$, $\rho(A_1) = 1$, $A_2 = [2]$, $\rho(A_2) = 2$, $A_3 = [1]$, $\rho(A_3) = 1$, $A_{12} = \left[\begin{smallmatrix} 1 & 2 \\ 0 & 2 \end{smallmatrix}\right]$, $\rho(A_{12}) = 2$, $A_{13} = \left[\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right]$, $\rho(A_{13}) = 2$, $A_{23} = \left[\begin{smallmatrix} 2 & 0 \\ 0 & 1 \end{smallmatrix}\right]$, $\rho(A_{23}) = 2$.

5. Существует $\alpha > 1$, такое что $0 \le A < \alpha A \le B$. Далее при $\rho(A) \ne 0$ использовать следствие 9.1, а при $\rho(A) = 0$ — результат задачи 6.4, примененный к матрице B.

6. $5 \le \rho(A) \le 9$.

7. Использовав следствие 9.3 для A^k при $\alpha = 0$, показать, что $\rho(A^k) > 0$. Затем доказать, что если λ_i, $i = \overline{1,n}$ – собственные значения матрицы A, то λ_i^k, $i = \overline{1,n}$ – собственные значения матрицы A^k.

9. $\rho(B) = \rho(A) = 0$. B не является неотрицательной. Доказыватся от противного с использованием задачи 6.4 стр. 187.

11. Непосредственно по определению перронова вектора.

12. Использовать неравенства из теоремы 9.3 стр. 309, записанные для положительного собственного вектора x.

14. Использовать упражнение 12 и утверждение c) теоремы Перрона.

15. Использовать нормировку перронова вектора.

16. Доказать и использовать тот факт, что если $A > 0$, $B \ge 0$ и B – невырожденная, то $AB > 0$.

17. Рассмотреть матрицы $\left[\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right]$, $\left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right]$ и $\left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]$.

18. Применить теорему Перрона к матрице A^k.

20. a) да, b) нет, так как является разложимой, c) неразложима, но импримитивна.

21. a) 1, c) 4.

22. a) да, b) нет.

23. Использовать определение унитарной матрицы.

24. Использовать определение двоякостохастической матрицы и теорему 9.17.

25. a) $L = \begin{bmatrix} \frac{3+\sqrt{5}}{2} & \frac{1}{2}(1+\sqrt{5}) \\ \frac{1}{2}(1+\sqrt{5}) & 1 \end{bmatrix}$, b) $L = \begin{bmatrix} \frac{1}{2} & \sqrt{\frac{1}{6}} \\ \sqrt{\frac{3}{8}} & \frac{1}{2} \end{bmatrix}$.

26. b) перронов корень $\rho(A) = 1 + \sqrt{6}$, перронов вектор $x = \left[\frac{\sqrt{6}}{3+\sqrt{6}};\ \frac{3}{3+\sqrt{6}}\right]$.

28. Использовать теоремы 9.17, 9.20.

29. $A^{\infty} = \frac{1}{89}\begin{bmatrix} 35 & 25 & 29 \\ 35 & 25 & 29 \\ 35 & 25 & 29 \end{bmatrix}$.

33. Использовать определение примитивности и теорему 9.13.

Перечень примеров

Рекомендуемая литература

[1] Беллман, Р. Введение в теорию матриц / Р.Беллман — М.: Наука, 1969. — 367 с.

[2] Мальцев, А.И. Основы линейной алгебры / А.И.Мальцев — М.: Наука, 1975. — 398 с.

[3] Ланкастер, П. Теория матриц / П.Ланкастер — М.: Наука, 1982. — 272 с.

[4] Беклемишев, Д.В. Дополнительные главы линейной алгебры / Д.В.Беклемишев — М.: Наука, 1983. — 336 с.

[5] Воеводин, В.В. Матрицы и вычисления / В.В.Воеводин, Ю.А.Кузнецов — М.: Наука, 1984. — 320 с.

[6] Проскуряков, И.В. Сборник задач по линейной алгебре / И.В.Проскуряков — М.: Наука, 1984. — 336 с.

[7] Размыслович, Г.П. Геометрия и алгебра: учеб. пособие / Г.П.Размыслович, М.М.Феденя, В.М.Ширяев.— Минск.: Изд. «Университетское», 1987. — 352 с.

[8] Беклемишев, Д.В. Курс аналитической геометрии и линейной алгебры / Д.В.Беклемишев — М. : Наука, 1987. — 319 с.

[9] Гантмахер, Ф.Р. Теория матриц / Ф.Р.Гантмахер. — М.: Наука, 1988. — 552 с.

[10] Хорн, Р. Матричный анализ / Р.Хорн , Ч.Джонсон — М.: Мир, 1989.— 655 с.

[11] Магнус, Я.Р. Матричное дифференциальное исчисление с приложениями к статистике и эконометрике / Я.Р.Магнус, Х.Нейдеккер. — М.: Физматлит, 2002. — 496 с.

[12] Мазаник, С.А. Функции от матриц: учеб.-метод. пособие / С.А.Мазаник, Г.П.Размыслович, В.М.Ширяев.— Минск: БГУ, 2002. — 34 с.

[13] Маркус, М. Обзор по теории матриц и матричных неравенств / М.Маркус, Х.Минк. — М.: Едиториал УРСС, 2004. — 232 с.

[14] Цехан, О.Б. Матричный анализ. Часть 1. Учебное пособие / О.Б.Цехан. — Гродно: ГрГУ, 2004. — 78 с.

[15] Цехан, О.Б. Матричный анализ. Часть 2. Учебное пособие / О.Б.Цехан. — Гродно: ГрГУ, 2004. — 135 с.

[16] Задачи по матричному анализу / А.К.Деменчук [и др.]. — Минск.: БГУ, 2004.— 52 с.

[17] Комраков, Б.Б. Матричный анализ: курс лекций / Б.Б.Комраков. — Минск.: БГУ, 2006.-102 с.

[18] Буков, В.Н. Вложение систем. Аналитический подход к анализу и синтезу матричных систем / В.Н.Буков. — Калуга: Изд-во научной литературы Н.Ф. Бочкаревой, 2006.— 720 с.

[19] Матричный анализ в примерах и задачах: практикум для студентов фак. приклад. математики и информатики и мех.-мат. фак. /А.К.Деменчук [и др.] — Минск.: БГУ, 2008.— 158 с.

Предметный указатель

Учебное издание

Цехан Ольга Борисовна

МАТРИЧНЫЙ АНАЛИЗ

Учебное пособие

Редактор М.В.

Компьютерная верстка: О.Б.Цехан

Дизайн обложки:

Сдано в набор . Подписано в печать

Формат 60x84/16. Бумага офсетная №1.

Печать RISO. Гарнитура Таймс.

Усл. печ. л. Уч.-изд. л. Тираж 300 экз. Заказ

Учреждение образования “Гродненский государственный

университет имени Янки Купалы”.

ЛВ № 02330/01633257 от 30.04.2004. Гродно.

Отпечатано на технике издательского центра

Учреждения образования “Гродненский государственный

университет имени Янки Купалы”.

ЛВ № 02330/0056882 от 30.04.2004. Гродно.